

Durham Research Online

Deposited in DRO:

15 July 2014

Version of attached file:

Published Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Schaller, Matthieu and Becker, Claude and Ruchayskiy, Oleg and Boyarsky, Alexey and Shaposhnikov, Mikhail (2014) 'A new framework for numerical simulations of structure formation.', *Monthly notices of the Royal Astronomical Society.*, 442 (4). pp. 3073-3095.

Further information on publisher's website:

<http://dx.doi.org/10.1093/mnras/stu1069>

Publisher's copyright statement:

This article has been accepted for publication in *Monthly Notices of the Royal Astronomical Society* © 2014 The Authors Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.



A new framework for numerical simulations of structure formation

Matthieu Schaller,¹★ Claude Becker,² Oleg Ruchayskiy,² Alexey Boyarsky^{3,4}
and Mikhail Shaposhnikov²

¹*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*

²*Institut de Théorie des Phénomènes Physiques, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

³*Instituut-Lorentz for Theoretical Physics, Universiteit Leiden, Niels Bohrweg 2, Leiden, the Netherlands*

⁴*Bogolyubov Institute of Theoretical Physics, Kyiv, 03680 Ukraine*

Accepted 2014 May 29. Received 2014 May 21; in original form 2013 October 27

ABSTRACT

The diversity of structures in the Universe (from the smallest galaxies to the largest super-clusters) has formed under the pull of gravity from the tiny primordial perturbations that we see imprinted in the cosmic microwave background. A quantitative description of this process would require description of motion of zillions of dark matter particles. This impossible task is usually circumvented by *coarse graining* the problem: one either considers a Newtonian dynamics of ‘particles’ with macroscopically large masses *or* approximates the dark matter distribution with a continuous density field. There is no closed system of equations for the evolution of the matter density field alone and instead it should still be discretized at each time step. In this work, we describe a method of solving the full six-dimensional Vlasov–Poisson equation via a system of auxiliary Schrödinger-like equations. The complexity of the problem gets shifted into the choice of the number and shape of the initial wavefunctions that should only be specified at the beginning of the computation (we stress that these wavefunctions have nothing to do with quantum nature of the actual dark matter particles). We discuss different prescriptions to generate the initial wavefunctions from the initial conditions and demonstrate the validity of the technique on two simple test cases. This new simulation algorithm can in principle be used on an arbitrary distribution function, enabling the simulation of warm and hot dark matter structure formation scenarios.

Key words: methods: numerical – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

The Lambda cold dark matter (Λ CDM) cosmological model is the current theoretical framework to describe the formation and evolution of large-scale structures in the Universe. In this model, the growth of structures occurs through the hierarchical collapse of a collisionless fluid of cold dark matter (CDM). Small initial perturbations grow through merging to create more and more massive haloes and complex substructures (e.g. Davis et al. 1985; Bertschinger 1998; Springel et al. 2005). These initial perturbations are thought to be (almost) Gaussian, created from quantum fluctuations during the inflation epoch and are the origin of all the objects seen in the Universe. The knowledge of the precise initial conditions and a comprehensive understanding of the underlying physical laws should, in principle, enable us to evolve these fluctuations forward in time and provide a test of the current models.

Most of the important features observable in the Universe today have grown via non-linear evolution from tiny primordial density perturbations. This makes the whole process of understanding their evolution complex and requires the use of techniques well beyond the linear perturbation theory (Bernardeau et al. 2002). Indeed, at scales below roughly 10 Mpc the evolution of structures had already entered the non-linear stage (i.e. the density contrast $\delta\rho$ is of order (or much greater) than the background density $\bar{\rho}$). The main resource available to cosmologists is the use of bigger and bigger cosmological simulations, most of them using the particle technique known as N -body simulation (Hockney & Eastwood 1988; Dehnen & Read 2011). Numerical simulations may, for instance, help shed some light on the unknown nature of dark matter (DM).

Clearly, the number of DM particles is way too large to track individually each of them on a computer. Therefore, most of the cosmological N -body simulations use macroscopically large simulation ‘particles’ (with their masses ranging from masses much larger than DM particles up to the size of a small galaxy, 10^8 – $10^9 M_\odot$).

★ E-mail: matthieu.schaller@durham.ac.uk

The problem of DM evolution in the Universe can be formulated as an evolution of a collisionless self-gravitating fluid.

The main tool used to describe this DM fluid is the *phase space density distribution* $f(x, v, t)$, defined such that $f(x, v, t)dqdv$ represents the mass of material at position x moving at velocity v at time t . This function is usually normalized such that its integral over all positions and velocities gives the total mass

$$\int d^3x \int d^3v f(x, v, t) = M_{\text{tot}}. \quad (1)$$

Notice that one could also normalize this integral to one or to the total number of particles in the system. When integrating over velocity space only, one gets the usual mass density $\rho(x)$, whereas integrating over all space returns the velocity distribution $d_v(v)$:

$$\rho(x) = \int f(x, v, t) d^3v, \quad d_v(v) = \int f(x, v, t) d^3x. \quad (2)$$

This distribution function obeys the Liouville theorem (Binney & Tremaine 2008) and if the only force acting on the particle is the gravitational potential $U(x)$, we can write a closed system of equations for the formation of structures (Bertschinger 1995; Bernardeau et al. 2002):

$$\frac{\partial f}{\partial \tau} + \frac{v}{a(\tau)} \frac{\partial f}{\partial x} - a(\tau) \nabla U \frac{\partial f}{\partial v} = 0, \quad \nabla^2 U = 4\pi G a^2(\tau) \delta \rho, \quad (3)$$

where x and v are comoving coordinates and velocities, $a(\tau)$ is the scalefactor and τ is the conformal time. (We will use this convention throughout this paper.) This Vlasov–Poisson (V-P) system has no solution in the general case and the only way to handle it is to use numerical techniques. For completeness, we also give the expressions for the density and density contrast:

$$\rho(x, \tau) = \frac{1}{a^3(\tau)} \int f(x, v, \tau) d^3v, \quad (4)$$

$$\delta \rho(x, \tau) = \frac{1}{a^3(\tau)} \left(\int f(x, v, \tau) d^3v - \frac{M_{\text{tot}}}{V_{\text{tot}}} \right), \quad (5)$$

where V_{tot} is the total comoving volume over which we average.

2 STRUCTURE FORMATION SIMULATIONS

The numerical analysis of the V-P system of equations (3) is very challenging. The first reason is that the system is six dimensional. Recent simulations can only handle up to 64 resolution elements in each spatial and velocity space direction (Yoshikawa, Yoshida & Umemura 2013) due to memory restrictions. Even the use of the biggest supercomputers would not allow us to go much beyond this figure.

The second shortcoming of such technique is the development of fine-grained structures that are very difficult to follow numerically. These become very important in structure formation scenarios as clusters typically present many matter streams and shell crossings.

Those two main shortcomings make the search for more advanced numerical scheme important. The problem of high dimensionality could be removed if there were a way to use the density field $\rho(x)$ instead of the probability distribution function $f(x, v)$. This can be done by integrating the first few moments of the Vlasov equation and then use techniques known for hydrodynamical simulations (see e.g. Hockney & Eastwood 1988). This technique is limited by the formal need to integrate all moments and not just the first few

ones to obtain an exact solution. Instead of a six-dimensional space, there is now a (formally) infinite number of variables obeying an infinite series of equations. Peebles (1987), for instance, truncates the series and uses the first two moments (mass conservation, Euler equation) of the collisionless Boltzmann equation to evolve in time the initial perturbations. The framework reaches its limits whenever the velocity dispersion of the fluid becomes important or when shell crossing occurs.

2.1 *N*-body simulations

The other option to solve the system of equations (3) is to use a particle method in which the distribution function is sampled by a finite number N of particles such that

$$f(x, v) \cong \frac{1}{N} \sum_{i=1}^N m_i \delta(x - x_i) \delta(v - v_i). \quad (6)$$

Each particle or body is then evolved according to Newton's law under the influence of the gravitational potential created by all the others as described by Poisson's equation. In other words, *N*-body simulations solve the Vlasov equation via its characteristics by sampling the initial phase space distribution with a discrete number of particles. The number of bodies is typically chosen as large as computationally feasible. The *N*-body formalism is thus a Monte Carlo approximation of the V-P system. The advent of large supercomputers combined with the development of more efficient numerical algorithms has enabled the field of cosmological simulations to make considerable progress over the last decades. Simulations such as the *Millennium run* (Springel et al. 2005) or *Bolshoi simulation* (Klypin, Trujillo-Gomez & Primack 2011) are able to follow as many as a few billion particles.

The complicated part of the *N*-body simulation is the evaluation of the forces between pairs of particles. Over the years, many ingenious techniques (see Dehnen & Read 2011 for a review) have been invented to reduce the algorithms complexity for the force integration to $\mathcal{O}(N \log N)$ or even better (Dehnen 2000). All these techniques (tree-code, particle-mesh, P³M, AMR, tree-PM, ...) do however rely on particles and do, hence, share the same initial assumptions leading to the two following challenges.

First, since the DM fluid is supposed to be collisionless, one has to manually suppress artificial two-body collisions arising between the pseudo-particles introduced to sample the phase space distribution. This is usually done by introducing an ad hoc softening length and suppressing the gravitational force at scales below it (Dehnen 2001). *N*-body simulations are run under the assumption that for a suitable choice of the smoothing, the evolution of the *N* pseudo-particles under the softened force should be the same as the gravitational evolution of the elementary DM particles.

The second challenge is to relate the particle distribution to the theoretical V-P the particles are supposed to model. Despite its obvious relevance, it seems that the question of the precise quantitative importance of the discretization (6) and its effects is still not settled (Joyce 2008).

As a matter of fact, there are no alternative tools to study the cosmic structure formation with the same resolution as *N*-body simulations. This is of course not a limitation of the *N*-body method itself, but makes it more complicated to evaluate the possible errors of *N*-body simulations quantitatively, as there are basically no independent results to compare with. For instance, Ludlow & Porciani (2011) find a non-negligible fraction of haloes in CDM simulations that cannot be matched to peaks in the initial density distribution

and are possible artefacts of the N -body method. The different techniques used to calculate the forces are, of course, different and can lead to marginally different results for the same initial sampling of the field when the resolution limit is reached. They do, however, all share the decomposition of $f(x, p)$ in a set of N macroscopic particles and will, hence, share the consequences of this Ansatz.

Spurious effects due to the discretization become more apparent when looking at simulations of warm dark matter (WDM) or hot dark matter (HDM) cosmologies. The initial matter power spectrum entering such simulations is truncated below a certain free-streaming scale related to the DM particle rest mass. Those particles having a small mass, they also have a finite velocity distribution function at every point in space, making the problem effectively six dimensional. In practice, these velocities are neglected and the DM fluid is treated in the cold fluid limit. These simulations are run using the same N -body framework but with an initial density and velocity power spectrum truncated below the scale of interest. This should lead to a suppression of small haloes below a characteristic mass and the simulations ought to be able to reproduce all structures with a mass above this limit. They could thus quickly converge towards a solution. Colín, Avila-Reese & Valenzuela (2000), Wang & White (2007) and Colín, Valenzuela & Avila-Reese (2008) did, however, demonstrate that this is not the case and that spurious haloes form and merge to form structures below the theoretical mass threshold. Various techniques are used in the literature to cure this problem. Lovell et al. (2012), for instance, filter their halo catalogues during the post-processing of their simulations. The end results are thus free from spurious haloes but it does not solve the intrinsic discreteness problem of the N -body technique.

More details about these challenges and a comprehensive review of the topic can be found in Dehnen & Read (2011). Notice that this formalism is still a very active and lively area of research with alternative more advanced formulations being proposed frequently. Some authors (Abel, Hahn & Kaehler 2012; Shandarin, Habib & Heitmann 2012) proposed recently to use tessellations of the three-dimensional matter sheet in six-dimensional space to track some of the phase space information. This may allow them to solve the coarse graining problem and reduce the impact of non-physical two body relaxations between the macroscopical particles. This formalism has led to promising results in the study of WDM cosmology and the differences between the CDM and WDM halo mass functions (Angulo, Hahn & Abel 2013).

All the potential shortcomings of the N -body formalism and the difficulty to evaluate their impact on the simulation results make it important to develop another framework not based on a particle approach.

2.2 An alternative framework

Our framework resembles the attempt by Peebles (1987) to use only the density field $\rho(x)$ and potential $U(x)$. The main problem of such an approach is that there is no closed system of equations that includes only the density and gravitational potential.

The situation is different when looking at quantum physics. In this realm, all the phase space information can be encoded in a single function, the wavefunction $\psi(x)$ which does not depend on the velocity v . It is thus possible to write a closed Schrödinger–Poisson system that would replace the V-P one and that would only depend on the spatial variable x . (See also Short & Coles 2006 for a similar idea.) This would effectively be a three-dimensional system of equations but would allow to simulate the full six-dimensional phase space and hence allow simulation of alternative cosmologies,

such as the one including WDM or free-streaming neutrino contributions. The principal difficulty is then to find a good mapping between the distribution function $f(x, v)$ of interest and its ‘quantum’ equivalent $\psi(x)$ and vice versa. This is achieved by using the so-called *Wigner distribution function* (WDF)

$$f(x, v) \simeq \int e^{\frac{i}{\hbar}vy} \psi^* \left(x + \frac{y}{2} \right) \psi \left(x - \frac{y}{2} \right) d^3y. \quad (7)$$

which obeys an equation similar to the Vlasov equation but is constructed from wavefunctions. The main feature of this mapping is that the density field can simply be expressed as

$$\rho(x) = |\psi(x)|^2. \quad (8)$$

However, the limitation of this approach is that one single wavefunction is in general not sufficient to encode all the complexity of the distribution function and we would then use the more general version:

$$\rho(x) = \sum_n |\psi_n(x)|^2. \quad (9)$$

The summation index n can, as a first thought, be understood as a sum over the velocities v that appear in the distribution function $f(x, v)$. We somehow trade a six-dimensional function for a (finite) set of three-dimensional (complex-valued) functions. We will, however, demonstrate that the number of wavefunctions required can be very low (of the order of unity in some cases), making the whole framework effectively three dimensional.

It is important to stress from the onset that we are not trying to solve the evolution of structure formation at the quantum level. Although we make use of quantum mechanics concepts, we merely use it as mathematical ‘trick’ to solve the V-P system (3). For this reason, the constant \hbar appearing in our equations has to be understood as a computational parameter whose value bears no relation to the actual Planck constant $\hbar_{\text{phys}} = 1.0545 \cdot 10^{-34} \text{ m}^2 \text{ kg s}^{-1}$.

Once the wavefunctions are built, they are evolved forward in time using Schrödinger’s equation. The density sourcing Poisson’s equation is obtained through equation (9) and one can then solve for the potential at each time step using standard techniques. This potential enters the Schrödinger equation, closing the loop. We have, hence, built a closed system of equations using only a set of wavefunctions (which serve as a proxy for density) and the gravitational potential:

$$i\hbar\partial_t\psi_n = -\frac{\hbar^2}{2}\nabla_x^2\psi_n + U\psi_n, \quad \nabla_x^2U = 4\pi G\delta\rho. \quad (10)$$

The study of structure formation then becomes an exercise in solving n copies of the Schrödinger equation on a computer, which is a well-studied problem. The velocity distribution can be recovered by Fourier transforming the wavefunctions and if one is interested in the phase space distribution, one can apply the Wigner transform. This is, however, not part of the algorithm itself. This can be done in post-processing if necessary. The entire evolution of the system can be done at the ‘quantum level’, i.e. using the wavefunctions alone.

We stress that this is another approximation of the true underlying physical problem (equation 3) and that this framework, as any other, will have limitations. Some of these limitations and their relevance to the case of structure formation studies will be discussed in this paper. We will address those in the context of the science we are interested in and demonstrate how alternative cosmologies, including non-CDM scenarios, could effectively be simulated.

The development of this framework has been pioneered by Widrow & Kaiser (1993) and Davies & Widrow (1997) with the important difference that these authors use a single wavefunction and another way to map the distribution function in the quantum world. Their general procedure is very similar to ours: sample the wavefunction from the initial phase space distribution, evolve in time using the Schrödinger–Poisson equations and recover the final phase space distribution from the wavefunction.

Note also that another possible route, where the Hartree equation is used instead of the Schrödinger equation, has been explored by Aschbacher (2001) and Fröhlich, Tsai & Yau (2010).

The time evolution of the Schrödinger–Poisson system is done using an explicit finite differences scheme for the wavefunction and a fast Fourier transform (FFT) algorithm to solve the Poisson equation. We try to improve upon their algorithm for the time evolution as will be described below. Widrow & Kaiser (1993) have made several simulations using this Schrödinger method obtaining results in agreement with usual N -body simulations. They also claim that their method is computationally comparable to N -body simulations making it a promising tool for cosmological purposes.

These authors choose to use one single wavefunction to represent the distribution function. This has important consequences on the validity of equation (8). By using one single wavefunction, the phase space distribution built from it cannot be everywhere positive and the authors have thus to add an additional Gaussian smoothing. We alleviate this shortcoming by using more than one wavefunction and a different transformation from wavefunctions to phase space distribution.

Their choice of Gaussian smoothed density also led them to a simple technique to generate the initial wavefunction. They use a set of particles sampling the phase space distribution function exactly as in the case of N -body simulations. They can then turn each particle into a Gaussian in phase space by smoothing it and use this set of wavepackets as their initial wavefunction.

In our approach, we depart from this need of an initial N -body sampling by considering other techniques to generate the set of wavefunctions. By doing so, we allow for a completely generic distribution function and should, in principle, not experience the consequences of an a priori artificial Monte Carlo sampling of $f(x, v)$.

The second feature of our framework is the replacement of the Poisson equation by a Klein–Gordon equation for the potential $U(x)$:

$$-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} U + \nabla_x^2 U = 4\pi G \delta \rho, \quad (11)$$

where c is the numerical speed of gravity. This scalar gravity equation is, once again, purely a mathematical trick to reduce the complexity of the original system (3) and not an attempt to modify Newton’s gravity. Such a replacement makes the framework entirely local and does not require complicated integration methods for the Poisson equation. The complexity of the scheme is then formally reduced to $\mathcal{O}(M)$, where M is the number of mesh points in real space used in the simulation. In this respect, our approach also differs from the original work by Widrow & Kaiser (1993), who stick to the classical Poisson form of gravity.

We stress that this step is not formally necessary. The well-known techniques used to solve Poisson’s equation on a mesh (FFT, Gauss–Seidel relaxation, etc.) can also be used in our framework. This change of equation for gravity does just make the computations slightly faster in the cases where our approximation is valid. However, in the case of cosmological simulations with vastly different

scales interacting, it is unclear how the Klein–Gordon equation for gravity would behave and defaulting to standard mesh techniques might be required.

3 THE ALGORITHM IN BRIEF

Here we present the main algorithm of our framework, decomposed in a few simple steps. A formal derivation and a discussion of the convergence and accuracy of the method will be presented in the next section.

Step 0. Choose the parameters of your simulation. The precision and speed of the method is governed by three parameters \hbar , c and N . The algorithm of choosing them is the following.

The parameters c and \hbar are linked to the time and space resolution (Δx and $\Delta \tau$, respectively) of the simulation via the Courant condition:

$$\frac{\Delta x}{\Delta \tau} > c \quad (12)$$

and the condition on the stability of the discretized Schrödinger equation:

$$\frac{\Delta x^2}{\Delta \tau} > \hbar. \quad (13)$$

The number of wavefunctions N is chosen depending on the number of relevant modes of the decomposition in wavefunctions of the initial distribution function. The optimal value of N is problem dependent and is also influenced by the algorithm chosen to discretize the distribution function. The details of this procedure will be given in Section 5. The precision of the original accuracy is also dictated by the choice of \hbar . The ‘quantum’ nature of the formalism imposes limitations on the precision of the description of position and velocity at the same point following the equivalent of Heisenberg’s uncertainty principle.

Step 1. Take an initial phase space distribution function of the matter fields $f(x, v)$ (in the case of cosmological simulations, it is expressed via the power spectrum $P(k)$). Decompose the distribution function in N complex-valued $\psi_n(x)$ such that

$$f(x, v) \approx \sum_{n=1}^N \int e^{\frac{i}{\hbar} v y} \psi_n^* \left(x + \frac{y}{2} \right) \psi_n \left(x - \frac{y}{2} \right) d^3 y. \quad (14)$$

The number N of wavefunctions is chosen such as to minimize the error introduced by the decomposition and will, in practice, be as big as computationally feasible. Various ways to generate this initial set of wavefunctions for a given $f(x, v)$ are presented in Section 5.

At this stage the precision of the approximation $f(x, v) \rightarrow \{\psi_n(x)\}$ is controlled by two parameters, \hbar and N .

Step 2. The wavefunctions $\psi_n(x)$ are now evolved forward in time using the coupled Schrödinger–Klein–Gordon system of equations

$$\begin{aligned} i\hbar \partial_t \psi_n &= -\frac{\hbar^2}{2} \nabla_x^2 \psi_n + U \psi_n, \\ -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} U + \nabla_x^2 U &= 4\pi G \left(\sum_{n=1}^N |\psi_n(x)|^2 - \bar{\rho} \right). \end{aligned} \quad (15)$$

The integration in time of the Schrödinger–Klein–Gordon system can be done explicitly using finite differences on a regular grid, as will be described in Section 6.

Step 3. Controlling your simulation. As the simulation is running you should monitor the following quantities in order to see that the

choice of the method does not introduce artefacts. The correction terms

$$\sum_{r \geq 3} \frac{1}{r!} \left(\frac{\hbar}{2i} \right)^{r-1} \frac{\partial^r}{\partial x^r} V \frac{\partial^r}{\partial v^r} f(x, v) \quad (16)$$

should be small when compared to the ones ($v \frac{\partial f}{\partial x}$ and $\frac{\partial U}{\partial x} \frac{\partial f}{\partial v}$) entering the Vlasov equation. Thanks to the $1/r!$ decrease and the smoothness of the gravitational potential V in most cases of interest, computing the first term of this series is generally sufficient. If this term grows above the value of the other terms in the Vlasov equation, then the approximation introduced in this paper is not valid any more. Reducing the value of \hbar or increasing the number N of wavefunctions used in the initial discretization will decrease the contribution of the correction terms but this will lead to a higher computational cost. The correction terms as well as the terms entering the Vlasov equation are expensive to compute but need not be computed at each time step.

Step 4. Once the final time has been reached, the distribution function can be recovered by computing the integral (14) or if one is only interested in the density field only, then equation (9) is sufficient and straightforward to compute.

4 FORMAL DERIVATION

In the previous section, we described the problem we were interested in and the usual schemes used in the literature. We also presented a brief description of the route we intend to follow in order to tackle the issues outlined. In this section, we describe the whole formulation in detail, derive its main equations and discuss its limits. For completeness, we start with a review of a formulation of quantum mechanics and show how its main ingredient, the WDF, will play the role of an approximate distribution function for our problem. Readers interested only in the end results can jump directly to Section 4.5.

4.1 Phase space quantum mechanics

Quantum mechanics is usually presented as emerging from the Hamiltonian formulation of classical mechanics through canonical quantization. (See for instance Sakurai & Napolitano 2011.) In this procedure, variables are promoted to Hermitian operators and the Poisson bracket is replaced by a commutator. Alternatively, one can also use Feynman's propagator and the path integral formalism to move from classical to quantum mechanics.

Alongside these well-known quantization procedures, there exist other equivalent formulations which try to emphasize more clearly certain aspects. The Moyal (or phase space) formulation is among those and tries to find a quantum equivalent to the classical phase space and distribution functions (Hillery et al. 1984; Ercolessi et al. 2007). The quantization procedure tries to find a correspondence between classical functions (called symbols) of the phase space variables and quantum operators in Hilbert space:

Operators in Hilbert space \leftrightarrow Phase-space symbols. (17)

As the position and momentum operators do not commute, this mapping cannot be unique. Different operator orderings will be mapped to different phase space symbols. Hermann Weyl proposed a systematic way to associate a quantum operator to a classical distribution function, which is now referred to as *Weyl quantization*. This complex procedure will not be discussed further here but its inverse, the *Wigner transform* will be useful for our formalism. This

transformation associates to every quantum operator \hat{A} a real phase space function $A(x, v)$:

$$A(x, v) = \text{sym}(\hat{A}) := \int e^{\frac{i}{\hbar}vy} \left\langle x - \frac{y}{2} \right| \hat{A} \left| x + \frac{y}{2} \right\rangle dy, \quad (18)$$

where $\langle \cdot | \cdot \rangle$ is the usual Bracket notation for quantum states. The transformation of a product of operators is given by

$$\text{sym}(\hat{A}\hat{B}) := \text{sym}(\hat{A}) \star \text{sym}(\hat{B}), \quad (19)$$

where the *Moyal star product* \star contains the quantum mixing of the operators. This product of functions in phase space is defined as

$$A(x, v) \star B(x, v) := A(x, v) e^{\frac{i\hbar}{2} (\overleftarrow{\partial}_x \overrightarrow{\partial}_v - \overrightarrow{\partial}_v \overleftarrow{\partial}_x)} B(x, v) \quad (20)$$

and is a central element in this formulation of quantum mechanics. Defining the *Moyal bracket* (Moyal 1949) by

$$\{A, B\}_M := A \star B - B \star A \quad (21)$$

the commutator of operators is associated with the Moyal brackets of two symbols in the following way:

$$\text{sym}([\hat{A}, \hat{B}]) = \{ \text{sym}(\hat{A}), \text{sym}(\hat{B}) \}_M. \quad (22)$$

The dynamical equation in this formulation of quantum mechanics can be written in a simple way using these brackets and reads

$$i\hbar \partial_t f = \{ \hat{H}, f \}_M, \quad (23)$$

where \hat{H} is the Hamiltonian of the system. The interesting property of this formulation of quantum mechanics is that in the semiclassical limit $\hbar \rightarrow 0$, the dynamical equation reduces to the classical equation of motion expressed in terms of Poisson brackets

$$\partial_t f = \{H, f\}_P = H \left(\overleftarrow{\partial}_x \overrightarrow{\partial}_v - \overrightarrow{\partial}_v \overleftarrow{\partial}_x \right) f. \quad (24)$$

This illustrates how the algebraic structures of classical and quantum mechanics are related through the continuous changing of the parameter \hbar . This is the reason why such an approach to quantum mechanics is known as *deformation quantization* (Hirshfeld & Henselder 2002).

Let us now stop this overview and move to the part of this formalism which will be useful for the construction of our new simulation framework.

4.2 Wigner distribution function

The Wigner transform (equation 18) maps a quantum operator \hat{A} to a classical function in phase space. Wigner used this to associate a real phase space function to a quantum system (Wigner 1932), now called the WDF. It is defined as the symbol in phase space associated with the density operator $\hat{\rho}$:

$$P_W(x, v) := \text{sym}(\hat{\rho}) = \int e^{\frac{i}{\hbar}vy} \left\langle x - \frac{y}{2} \right| \hat{\rho} \left| x + \frac{y}{2} \right\rangle d^3y. \quad (25)$$

As usual, the density operator can be expressed as the combination of pure state wavefunctions ψ_n :

$$\hat{\rho} = \sum_n \lambda_n |\psi_n\rangle \langle \psi_n|, \quad \lambda_n \geq 0, \quad \sum_n \lambda_n = 1. \quad (26)$$

For mixed states, the WDF is thus

$$P_W(x, v) = \int e^{\frac{i}{\hbar}vy} \sum_n \lambda_n \psi_n^* \left(x + \frac{y}{2} \right) \psi_n \left(x - \frac{y}{2} \right) d^3y, \quad (27)$$

while for a pure state, it reads

$$P_W(x, v) = \int e^{\frac{i}{\hbar}vy} \psi^* \left(x + \frac{y}{2}\right) \psi \left(x - \frac{y}{2}\right) d^3y. \quad (28)$$

To simplify the expressions, we will use the notation $x_{\pm} = x \pm \frac{y}{2}$ and $\psi_{n\pm} = \psi_n(x_{\pm})$ in what follows. The WDF has many similarities to the classical distribution function: $P_W(x, v)$ is a real function, as can be seen by taking the conjugate and performing the change of variable $y \rightarrow -y$. It is also normalized to 1 in the following sense:

$$\int d^3x \int \frac{d^3v}{(2\pi\hbar)^3} P_W(x, v) = 1. \quad (29)$$

It has similar marginal distributions as can be seen by integrating over all velocities:

$$\begin{aligned} \int \frac{d^3v}{(2\pi\hbar)^3} P_W &= \sum_n \lambda_n \int d^3y \psi_n^*(x_+) \psi_n(x_-) d^3y \\ &= \sum_n \lambda_n |\psi_n(x)|^2, \end{aligned} \quad (30)$$

or over all space

$$\begin{aligned} \int d^3x P_W &= \sum_n \lambda_n \int d^3x_- d^3x_+ e^{\frac{i}{\hbar}vx_+} e^{-\frac{i}{\hbar}vx_-} \psi_{n+}^* \psi_{n-} \\ &= \sum_n \lambda_n \left| \tilde{\psi}_n \left(\frac{p}{\hbar}\right) \right|^2. \end{aligned} \quad (31)$$

In both cases the non-negative property of these marginal distributions is a property of the wavefunctions in quantum mechanics.

The WDF does, however, have the peculiar property that it may assume negative values. For this reason, it is called a *quasi-probability distribution* and cannot be interpreted as a phase space probability density in the sense of classical mechanics. The non-positivity of the WDF can be seen by integrating over all phase space the product of two distributions built from different states ψ and Φ :

$$\int dx \int dv P_W[\psi](x, v) P_W[\Phi](x, v) \propto |\langle \psi | \Phi \rangle|^2. \quad (32)$$

The right-hand side vanishes if the two states ψ, Φ are orthogonal which implies that the WDF cannot be positive everywhere. According to the Hudson theorem (Hudson 1974), the WDF of a pure state is point wise non-negative if and only if the state is Gaussian. If $\hat{\rho}$ is not a pure state, it can be represented as a convex combination of pure state operators, $\hat{\rho} = \sum_n \lambda_n |\psi_n\rangle\langle\psi_n|$, in infinitely many ways. The WDF satisfies the so-called mixture property (Ballentine 1998), which is the requirement that the phase space distribution should depend only on the density operator $\hat{\rho}$ and not on the particular way it is represented as a mixture of some set of pure states $\{|\psi_n\rangle\}$. To summarize, the WDF has many properties similar to the classical phase space distribution. Nevertheless, it has been realized from the early days, that the concept of a joint probability at a phase space point is limited in quantum mechanics because the Heisenberg uncertainty principle makes it impossible to simultaneously specify the position and velocity of a particle. Therefore, the best one can hope to do is to define a function that has a maximum of properties analogous to those of the classical distribution function. Many different variants of distribution functions – Husimi, Kirkwood–Rihaczek and Glauber – have been studied over the decades, all with their own advantages and shortcomings (see Lee 1995 for a review). The WDF is despite its non-positivity considered to be a useful calculational tool and finds applications in various domains outside of quantum physics, like signal processing or optics (Bastiaans 1997).

Widrow & Kaiser (1993) use a Husimi distribution (Husimi 1940) to recover the phase space information from the wavefunction. The Husimi distribution is essentially equal to the Wigner distribution with an additional Gaussian smoothing of width η

$$P_H(x, v) = \frac{1}{(2\pi\hbar)^3} \frac{1}{(\pi\eta^2)^{3/2}} \left| \int d^3y e^{-\frac{(x-y)^2}{2\eta^2} - \frac{i}{\hbar}vy} \psi(y) \right|^2. \quad (33)$$

Compared to the WDF it has the advantage of yielding a phase space distribution that is positive definite at every point. This comes at the price of the marginal distributions not being equal to the usual position and velocity distributions, but rather Gaussian broadened versions of it

$$\rho_H(x) = \frac{1}{(\pi\eta^2)^{3/2}} \int d^3y e^{-\frac{(x-y)^2}{2\eta^2}} |\psi(y)|^2. \quad (34)$$

Only in the limit $\eta \rightarrow 0$ does it reduce to the usual probability distribution. Similarly, one can show that the other marginal distribution reduces to the standard velocity distribution only when $\eta \rightarrow \infty$. This complementarity is of course related to Heisenberg's uncertainty principle. Note that it is in principle this smoothed distribution that enters Poisson equation instead of $|\psi(x)|^2$. Since this would be requiring an additional space integration at each time step, Widrow and Davies approximate it with the usual distribution $|\psi(x)|^2$ in the Poisson equation.

Actually, $P_H(x, p) \simeq f(x, p)$ only when averaged on scales $\Delta x \geq \eta$, $\Delta p \geq \frac{\hbar}{\eta}$. Note that there is no a priori reason why the non-linear time evolution should yield an answer that is again, in average, close to the real distribution function. Let us stress that we allow for several wavefunctions to have an initial phase space representation that is arbitrary close to the classical distribution function at every point, not only when averaged.

Let us recall that our goal is not to interpret the WDF as a fully fledged phase space distribution, but rather as a convenient mathematical tool.

4.3 Dynamical equation for the WDF

We now want to derive the dynamical equation satisfied by the WDF. A derivation starting from Liouville's equation for the density matrix can be found in Ballentine (1998). Another possibility is to start by taking the time derivative of the WDF and use the fact that the wavefunctions satisfy Schrödinger's equation.

Suppose each of the wavefunctions satisfies Schrödinger equation

$$i\hbar\partial_t \psi_n = -\frac{\hbar^2}{2} \nabla^2 \psi_n + V \psi_n, \quad (35)$$

then the time derivative of the WDF becomes

$$\begin{aligned} \partial_t P_W &= \int e^{\frac{i}{\hbar}vy} \sum_n \lambda_n \left[-\frac{i\hbar}{2} (\nabla_+^2 \psi_{n+}^* \psi_{n-} - \psi_{n+}^* \nabla_-^2 \psi_{n-}) - \frac{1}{i\hbar} (V_+ - V_-) \psi_{n+}^* \psi_{n-} \right] d^3y, \end{aligned} \quad (36)$$

where, once again, the subscripts $+, -$ denote the dependence on $x_{\pm} = x \pm \frac{y}{2}$. The terms containing a Laplacian can be rewritten in terms of spatial derivatives of P_W only and the previous equation becomes

$$\begin{aligned} 0 &= \partial_t P_W + \mathbf{v} \cdot \nabla_x P_W \\ &\quad - \frac{1}{i\hbar} \int e^{\frac{i}{\hbar}vy} (V_+ - V_-) \sum_n \lambda_n \psi_{n+}^* \psi_{n-} d^3y. \end{aligned} \quad (37)$$

This is the dynamical equation for the WDF, that we will refer to as the *Wigner equation*. This dynamical equation depends on both P_W and the wavefunctions which implies that we might have to define initial conditions for both. Let us now demonstrate that one can get rid of the dependency on the ψ_n . Let us expand the potential in Taylor series

$$V(x_+) - V(x_-) = y \frac{\partial}{\partial x} V(x) + 2 \sum_{r \geq 3, r \text{ odd}} \frac{1}{r!} \frac{\partial^r}{\partial x^r} V(x) \left(\frac{y}{2}\right)^2 \quad (38)$$

and use this result in the dynamical equation

$$0 = \frac{\partial}{\partial t} P_W + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{x}} P_W - \frac{\partial V}{\partial \mathbf{x}} \frac{\partial}{\partial \mathbf{v}} P_W + \sum_{r \geq 3, r \text{ odd}} \frac{1}{r!} \left(\frac{\hbar}{2i}\right)^{r-1} \frac{\partial^r}{\partial \mathbf{x}^r} V \frac{\partial^r}{\partial \mathbf{v}^r} P_W. \quad (39)$$

One can notice that the first three terms correspond to the classical Vlasov equation.

In three cases, the Wigner equation exactly coincides with the classical Vlasov equation: for a free particle ($V = 0$), for a uniform field ($V \propto x$) and for a harmonic oscillator ($V \propto x^2$). In general, there are additional terms that can be interpreted as quantum corrections¹ or simply higher order corrections. In any other case, corrections in the form of a power series in \hbar will appear and modify the dynamic. Note that in this derivation, the only assumption made on the λ_n is that they be constant. In principle, any value is acceptable and it can even be negative or complex. As we are not using these equations to solve a quantum mechanics problem, where $\lambda_n > 0$, we can use this fact to create more general sets of wavefunctions to approximate a given $f(x, v)$.

Note that the mass m does not appear in the Schrödinger equation in the same way as it does not appear in the Vlasov system. This, once again, illustrates that we are not solving the quantum mechanics evolution of the individual DM particles but rather find an approximation of the DM fluid evolution equation.

Let us recap what we have derived so far. By inspecting the Moyal formulation of quantum mechanics, we found a quantity, the WDF P_W . This quasi-probability density function obeys the Wigner equation, an equation similar to the Vlasov equation but with additional terms in the form of a power series in \hbar .

4.4 Semiclassical limit

The Wigner equation (39) reduces to the classical Vlasov equation in the limit $\hbar \rightarrow 0$. Even though the quantum correction is formally $\mathcal{O}(\hbar^2)$, the derivatives of P_W could generate additional inverse powers of \hbar , making the semi-classical limit more involved.²

The properties of the semiclassical limit depend of course on the potential $V(x)$. In this paragraph, we present some results concerning the case of interest to us, where the potential satisfies Poisson's equation. In particular, different authors investigated the semiclassical limit of the Wigner–Poisson (W-P) system to the V-P system for the Coulomb potential.

¹ It may sound surprising that the equation for the harmonic oscillator reduces exactly to the classical Vlasov equation, even though we know that the quantum mechanical treatment introduces discrete energy levels. In this case the quantum information is encoded purely in the initial conditions.

² This formulation of the statement is not fully satisfying, as the true semiclassical limit is also a statement about the properties of the wavefunction, and not identical to sending $\hbar \rightarrow 0$ which is anyway a dimensional parameter.

The mathematically rigorous classical limit from W-P to V-P has been solved first in 1993 independently by Lions & Paul (1993) and Markovitch & Mauser (1993). Both references consider a so-called completely mixed state; i.e. an infinite number of pure states with a strong additional constraint on the occupation probabilities:

$$\text{Tr} \hat{\rho}^2 = \sum_n \lambda_n^2 \leq C \hbar^3, \quad (40)$$

where C is a constant. Under this assumption, the classical limit of the solution to the three-dimensional W-P system converges to the solution of the V-P system. Note that the WDF can also have negative values, whereas the semiclassical limit is a true, non-negative distribution function. In both references, this was overcome by using a Gaussian-smoothed Wigner function.

The situation for a pure state is completely different (Zhang, Zheng & Mauser 2002). According to these authors, it appears that a density operator which has the above property that the trace of its square tends to zero with the third power of the Planck constant seems to be closer to classical mechanics than a pure state. For a pure state in one dimensional, the semiclassical limit is not unique: examples have been constructed where different regularization schemes give different limits (Majda, Majda & Zheng 1994). The question whether there exists a selection principle to pick the correct classical solution has also been investigated but is not yet settled (Jin, Liao & Yang 2008). No proof of the semiclassical limit from W-P to V-P is known for the pure state case in two dimensional or three dimensional.

For more details the reader is referred to the original papers or the review (Mauser 2002). See also Fröhlich, Graffi & Schwarz (2007) for an alternative approach to the semiclassical limit.

Finally, let us stress once again, that we seek to use our knowledge of quantum mechanics to simplify the resolution of the mathematical problem presented in Section 1. We are not trying to describe the physics of structure formation at the quantum level nor trying to find a wavefunction for the entire Universe.

4.5 Local interaction framework

In Newtonian gravity, much like in classical electrodynamics, each body moves in the potential generated by all the others. As both forces are long ranged, the total force acting on each of the N particles will be given by the sum of the contributions from all the other particles, no matter how far away. In gravitational N -body problems, the N sampling bodies also receive a contribution from all the other bodies and a naive algorithm would require $\mathcal{O}(N^2)$ operations for the force calculation at each time step. But it is well known that this long-ranged interaction through the potential can be replaced by a purely local interaction with a gauge boson or a spin-zero boson. In this approach, each particle only interacts locally with the bosonic field.

We propose to reformulate the cosmological V-P problem system (3)

$$\frac{\partial f}{\partial t} + \frac{v}{a} \frac{\partial f}{\partial x} - a \frac{\partial U}{\partial x} \frac{\partial f}{\partial v} = 0, \quad \nabla^2 U = 4\pi G a^2 \delta \rho, \quad (41)$$

as a purely local problem. To achieve spatial locality, we shall trade the real-valued phase space distribution function $f(x, v)$ for a finite set of complex-valued wavefunctions $\{\psi_n(x)\}$. For this, we shall assume that the classical distribution function can be approximated

by the WDF of some auxiliary mixed states:

$$f(x, v) \simeq P_W(x, v) = \int e^{\frac{i}{\hbar} v y} \sum_n \lambda_n \psi_n^*(x_+) \psi_n(x_-) d^3 y. \quad (42)$$

The details of how this approximation is to be understood, and how we construct in practice the set of wavefunctions $\{\psi_n(x)\}$ for any given $f(x, v)$ will be discussed in Section 5. For the time being, let us assume that we have determined a set of wavefunctions such that the above approximation holds.

The dynamical evolution of the WDF is given by the quantum-corrected Vlasov equation (the Wigner equation 39), or equivalently, by the Schrödinger equation (35) of the wavefunctions interacting in a self-consistent way with a potential obeying the Poisson equation. The cosmological Vlasov equation in an expanding universe and expressed using conformal time τ is very similar to the classical one, up to the replacements

$$v \mapsto \frac{v}{a(\tau)}, \quad V \mapsto a(\tau)U. \quad (43)$$

Therefore, the Schrödinger–Poisson system in the expanding universe becomes

$$i\hbar \partial_\tau \psi_n = -\frac{\hbar^2}{2a} \nabla_x^2 \psi_n + aU \psi_n, \quad (44)$$

$$\nabla_x^2 U = 4\pi G a^2 \delta\rho,$$

where $\delta\rho$ is the cosmological density contrast. The mass density [kg m^{-3}] relates to the wavefunctions [$\text{kg}^{1/2} \text{m}^{-3/2}$] by

$$\rho = \frac{1}{a^3} \int \frac{d^3 v}{(2\pi\hbar)^3} f(x, v) = \frac{1}{a^3} \sum_n \lambda_n |\psi_n|^2. \quad (45)$$

The normalization is chosen such that the phase space density integrates to the total mass

$$\int d^3 x \int \frac{d^3 v}{(2\pi\hbar)^3} f(x, v) = \int d^3 x \sum_n \lambda_n |\psi_n|^2 = M_{\text{tot}}, \quad (46)$$

implying for the background density

$$\bar{\rho} = \langle \rho \rangle = \frac{1}{V_{\text{tot}}} \frac{1}{a^3} \int d^3 x \sum_n \lambda_n |\psi_n|^2 = \frac{1}{a^3} \frac{M_{\text{tot}}}{V_{\text{tot}}}, \quad (47)$$

where V_{tot} denotes the total comoving volume. Therefore the density contrast $\delta\rho$ reads

$$\delta\rho = \frac{1}{a^3} \left(\sum_n \lambda_n |\psi_n|^2 - \frac{M_{\text{tot}}}{V_{\text{tot}}} \right). \quad (48)$$

In the semiclassical limit ($\hbar \rightarrow 0$), the Schrödinger–Poisson system (44) formally reduces to the original V-P system describing gravitational structure formation.

Notice that the total mass is conserved by construction as the normalization of the wavefunctions is a constant of motion of the Schrödinger equation.

So far, we achieved locality in the sense that our set of equations does not explicitly depend on the velocity variable v . We traded our six-dimensional phase space density function for a (possibly infinite) set of complex-valued functions that depend on the space coordinate x only. The numerical complexity of the problem has thus been drastically reduced as long as the number of wavefunctions remains small. Before addressing this question, let us go one step further and discuss the second equation of our system (44).

The Poisson equation is a non-local equation as the Laplacian operator couples the contributions from the whole space. This can,

however, be changed by replacing the Laplacian by a d'Alembertian operator. With this change, the Poisson equation becomes a Klein–Gordon equation and our transformed cosmological problem now reads

$$i\hbar \partial_\tau \psi_n = -\frac{\hbar^2}{2a} \nabla_x^2 \psi_n + aU \psi_n, \quad (49)$$

$$-\frac{1}{c^2} \partial_{\tau\tau}^2 U + \nabla_x^2 U = 4\pi G a^2 \delta\rho.$$

This system is entirely local, meaning that it can be numerically evolved in time on a grid by summing contributions of local sampling points only. If the contribution of the term $-\frac{1}{c^2} \partial_{\tau\tau}^2 U$ becomes small, then this system reduces to the Schrödinger–Poisson system discussed previously. This is in particular true in the non-relativistic limit $c \rightarrow \infty$. It is important to understand that the speed c does not necessarily have to take the value of the physical speed of light (or of gravity) $c_{\text{phys}} = 299\,792\,458 \text{ m s}^{-1}$. It must simply be understood as a parameter that we can use to approach the physical problem we are interested in (equation 3). As for \hbar , we are free to choose this parameter in a way that is convenient for our simulations, as long as we remain in the non-relativistic limit, meaning that the gravitational field U propagates much faster than the matter fields ψ_n .

Note, however, that using a non-infinite speed for the mediator of gravity in cosmological simulations may also be of some physical interest as the Poisson equation is, formally, only a weak-field approximation of the underlying Einstein equations from which a finite speed for the gravity emerges. Thus, modifying this parameter may also yield interesting physical results.

Let us summarize what we achieved so far. Using the formalism derived in the Sections 4.1 to 4.4, we have been able to construct a completely local system of equations (49) which in the non-relativistic classical limit $\hbar \rightarrow 0$, $c \rightarrow \infty$ reduces to the problem of cosmological structure formation. The probability density function can be computed at any time using the definition of the WDF (equation 27) but we stress that this operation is in general not necessary as one is usually interested in the evolution of the mass density (equation 30) only.

Let us finally say that replacing the Poisson equation by a scalar field is not strictly necessary as the algorithmic complexity of the problem has already been drastically reduced by the introduction of the WDF. Having a Schrödinger–Poisson system to solve instead of equation (3) is more accurate than our final system (49). It does, however, simplify a lot the numerical algorithms in some cases and does not seem to impact heavily the results as long as the parameter c is chosen wisely. The effect of this choice on the evolution of highly clustered matter fields found in the low-redshift Universe has not, however, been explored.

4.6 Lagrangian formulation

The system of equations (49) can be derived from a Lagrangian density using the Euler–Lagrange equations. We consider a real scalar field U interacting with the complex scalar matter fields ψ_n . The Lagrangian for this system reads

$$\mathcal{L} = \frac{1}{2c^2} \dot{U}^2 - \frac{1}{2} (\nabla U)^2 + \zeta \bar{\rho} U + \zeta \sum_n \lambda_n \left[\frac{i\hbar}{2} (\psi_n^* \dot{\psi}_n - \dot{\psi}_n^* \psi_n) - \frac{\hbar^2}{2a} \nabla \psi_n^* \cdot \nabla \psi_n - |\psi_n|^2 U \right]. \quad (50)$$

The equations of motion are found to be

$$i\hbar\dot{\psi}_n = -\frac{\hbar^2}{2a}\nabla^2\psi_n + U\psi_n, \quad (51)$$

$$-\frac{1}{c^2}\ddot{U} + \nabla^2 U = \zeta \left(\sum_n \lambda_n |\psi_n|^2 - \bar{\rho} \right), \quad (52)$$

which is the system we derived in the previous section if we set $\zeta = \frac{4\pi G}{a}$. The Hamiltonian density corresponding to the Lagrangian (50) is given by

$$\begin{aligned} \mathcal{H} = & \frac{1}{2c^2}\dot{U}^2 + \frac{1}{2}(\nabla U)^2 \\ & + \zeta \sum_n \lambda_n \frac{\hbar^2}{2a} |\nabla\psi_n|^2 \\ & + \zeta \left(\sum_n \lambda_n |\psi_n|^2 - \bar{\rho} \right) U, \end{aligned} \quad (53)$$

which has a positive definite kinetic energy term for the scalar potential, as expected from a well-behaved theory.

One can also decompose this Hamiltonian in its various energy components. Doing so allows us to control the impact of the dynamic term for the field U and consider it a valid approximation of the underlying V-P problem when its value is much lower than the other energy components. Together with the computation of the higher order terms of the Wigner equation (39), this measure of the impact of $c \neq \infty$ gives us a measure of the approximations we made and can thus help us assess the validity of the outcome of our simulations.

5 GENERATING INITIAL CONDITIONS

In the previous section, we showed how one can trade the Vlasov equation for the phase space distribution function for Schrödinger's equation for the wavefunctions, as this allows for the introduction of a scalar field as the mediator of the gravitational force. Of course we do not require the wavefunctions to have any intrinsic physical interpretation. We rather consider them, just like the WDF, as a mathematical tool and not as fundamental entities. Still we are faced with the problem of how to determine a set of wavefunctions such that their WDF corresponds to the initial classical phase space distribution.

One possible approach is to start from a set of N particles sampling the phase space distribution function and build Gaussians centred on each point with a certain width η

$$|\eta(x_i, v_i)\rangle \propto e^{-\frac{(x-x_i)^2}{2\eta^2} - \frac{i}{\hbar}v_i \cdot x}. \quad (54)$$

The wavefunction is then obtained from the incoherent superposition of these wavepackets for each particle

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^N e^{i\phi_i} |\eta(x_i, v_i)\rangle, \quad (55)$$

where $e^{i\phi_i}$ is a random phase. This sampling procedure relies on the assumption that each particle has a well-defined velocity. It is unclear how it could be generalized to the case of WDM, where the velocity dispersion is important. We remove the need for this assumption by allowing for several wavefunctions. At the same time this allows us to represent any initial phase space distribution without relying on N -body sampling. Such an approach was used

by Widrow and Kaiser and is well suited for Husimi distributions as they contain an extra Gaussian smoothing. We will, instead, try to work directly with the distribution function without sampling it in particles and hence taking the risk of facing the coarse graining and discreteness effects (Section 2.1) that we are trying to avoid in our framework.

Since the wavefunctions encode both, the position and velocity information, a single wavefunction (pure state) can in general not be sufficient to describe a generic $f(x, v)$. One should rather look for a set of wavefunctions (mixed state). The more wavefunctions we allow for, the more freedom we have and the more accurately the WDF should represent any given distribution. At the same time the total number of wavefunctions should be as small as possible because this will reduce the computational complexity of our numerical simulations.

Given the classical distribution function $f(x, v)$, we want to expand it using the WDF Ansatz

$$f(x, v) = \sum_{n=1}^N \lambda_n \int e^{\frac{i}{\hbar}vy} \psi_n^* \left(x + \frac{y}{2} \right) \psi_n \left(x - \frac{y}{2} \right) d^3y. \quad (56)$$

Fourier transforming from v -space to η -space we get

$$f(x, \eta) = \sum_{n=1}^N \lambda_n \psi_n^* \left(x + \frac{\eta}{2} \right) \psi_n \left(x - \frac{\eta}{2} \right). \quad (57)$$

Finding the wavefunctions is now a simpler problem provided one can easily compute the Fourier transform of the distribution function one is interested in. We will discuss different approaches to tackle this problem of determining the set of wavefunctions ψ_n and weights λ_n representing a given initial phase space distribution $f(x, v)$. Let us stress from the outset that these procedures need only to be used once at the beginning of a numerical simulation, to set up the initial conditions.

Last but not least, we need to emphasize that the number of wavefunctions is preserved by the quantum mechanical evolution. There is no evolution equation for λ_n . Only the shape of the ψ_n will change. This shows that it is the complexity of the initial conditions that dictates the number of wavefunctions required. In a setup where only a restricted number of harmonics are present in the initial probability distribution, already relatively few wavefunctions would be sufficient to represent the system and its time evolution.

5.1 Brute-force minimization

The first and obvious method we present to choose the initial wavefunctions is a brute-force minimization. The underlying idea is to define a functional measuring the total absolute error made by approximating the phase space distribution by the WDF Ansatz

$$\Phi := \int d^3q \int d^3\eta \left| f(x, \eta) - \sum_{n=1}^N \lambda_n \psi_{n+} \psi_{n-} \right|^2, \quad (58)$$

where, once again, $\psi_{n\pm} = \psi_n(x \pm \frac{\eta}{2})$. We can then determine a set of wavefunctions that minimizes this error. In practice, the minimization is most easily done via discretization on a lattice. The problem is then cast into a minimization of the scalar error function with a large number of variables corresponding to the values of the wavefunctions at the lattice points. For different $N = 1, 2, \dots$, we can determine the set of wavefunctions ψ_n and corresponding weights λ_n which minimizes the error. One can then compare the results for different N to find an optimal approximation with a high enough accuracy and a minimal number of wavefunctions.

Since we are not seeking a true quantum mechanical interpretation, let us consider the most general case of complex-valued weights. A naive minimization will not yield wavefunctions normalized to unity. Instead of adding this normalization as a constraint to the minimization, we remove the amplitude of the complex weights λ_n , and only keep their phases $e^{i\phi_n}$. The amplitudes of the weights are taken to be the norm of the wavefunctions, thereby normalizing them to unity. If we simply minimize the error functional, we will in general obtain wavefunctions that are not smooth enough on the lattice to be evolved numerically. For this purpose, it is useful to add a term of the form of a kinetic term to the functional that will allow us to enforce a certain degree of smoothness. We construct the kinetic term from the square of the derivatives with a certain overall factor χ to tune the smoothness:

$$\mathcal{K} = \chi \int d^3x \sum_{n=1}^N \left| \frac{\partial \psi_n}{\partial x} \right|^2. \quad (59)$$

Finally, we minimize this kinetic term with the total error summed over all lattice points

$$\mathcal{E} = \mathcal{K} + \Phi. \quad (60)$$

We have applied the method to cosmic initial conditions of CDM in the Zel'dovich approximation, for simplicity in a one-dimensional case. The results confirm the expectation that, increasing the number of wavefunctions, the total error is reduced. In the case we studied, it turned out that already a relatively small number of wavefunctions (compared for instance to the number of lattice points) was enough to achieve a reasonable accuracy. As usual with minimization procedures, there is no guarantee that the algorithm converges to a global minimum. This would for instance mean that one has to repeat the minimization with different initial random seeds and compare their outcomes. Also, even though this minimization was shown to work for a given phase space distribution $f(x, v)$, in practice it becomes computationally challenging even for rather small three-dimensional lattice sizes, as the number of variables in the minimization procedure grows quickly. Despite its applicability to any distribution function, the brute-force minimization might not be the best method to determine the initial wavefunctions.

5.2 Eigenvalue problem for Hermitian operator

We now turn our attention to obtaining an analytic solution to the problem of determining the initial wavefunctions. More precisely we will show how the Wigner Ansatz can be reformulated as an eigenvalue problem, which we can then solve analytically in some specific cases.

Since $f(x, \eta)$ is the Fourier transform of a real function $f(x, v)$, it satisfies the condition $f^*(x, -\eta) = f(x, \eta)$. Introducing the coordinates $x_{\pm} := x \pm \frac{\eta}{2}$, we can define

$$g(x_-, x_+) := f\left(\frac{x_+ + x_-}{2}, x_+ - x_-\right), \quad (61)$$

which is then Hermitian

$$g^*(x_+, x_-) = g(x_-, x_+). \quad (62)$$

Hilbert–Schmidt’s theorem states that any square-integrable Hermitian kernel can be expressed in terms of its spectral decomposition

$$g(x_-, x_+) = \sum_n \lambda_n \psi_n^*(x_-) \psi_n(x_+), \quad (63)$$

where the λ_n are real eigenvalues and ψ_n the set of orthonormal eigenfunctions with respect to the standard scalar product on $L^2(\mathbb{C}^3)$

$$\langle \psi_n | \psi_m \rangle := \int \psi_n^*(x) \psi_m(x) d^3x = \delta_{nm}. \quad (64)$$

The Fourier space WDF (equation 57) has exactly the same form as the spectral decomposition (equation 63). Therefore, we conclude that *any* given phase space distribution function $f(x, v)$ can be written exactly as a WDF, if need be with an infinite number of wavefunctions. The wavefunctions are the eigenfunction of the Hermitian operator $g(x_-, x_+)$ and its real eigenvalues correspond to the weights of the wavefunctions in the mixed state. Notice though, that they can in general take negative values, implying that we cannot give a full quantum-mechanical interpretation to the mixed state, as the corresponding density operator is not positive definite. Let us emphasize once more that we consider the wavefunctions as a mere mathematical tool.

Multiplying both sides of (63) by $\psi_\alpha(x_-)$ and integrating over x_- , the orthonormality of the eigenfunctions implies the following integral equation

$$\int g(x_-, x_+) \psi_\alpha(x_-) d^3x = \lambda_\alpha \psi_\alpha(x_+). \quad (65)$$

This equation shows that the determination of the wavefunctions reduces to finding the eigenfunctions of the Hermitian kernel g . Unfortunately, for a completely general phase space distribution function, the above equation might not allow an analytic solution.

This procedure can be generalized by allowing for a more general scalar product containing a non-trivial weight function $w(x)$:

$$\langle \psi | \phi \rangle_w := \int \psi^*(x) \phi(x) w(x) d^3x. \quad (66)$$

For such a scalar product, the eigenvalue decomposition of $g(x_-, x_+)$ still exists but the eigenfunctions are now orthonormal with respect to the weighed scalar product. The eigenvalue problem thus reads

$$\int g(x_-, x_+) \psi_\alpha(x_-) w(x_-) d^3x = \lambda_\alpha \psi_\alpha(x_+). \quad (67)$$

Let us emphasize that the weighted scalar product is only used to determine the wavefunctions whose WDF equals the classical distribution function. The choice of $w(x)$ is completely arbitrary and does not affect the properties of the WDF or the Schrödinger evolution of the wavefunctions. Clearly, the spectrum will depend on the choice of weight function. The additional freedom of choosing $w(x)$ could allow us to reduce the number of wavefunctions needed in the Wigner Ansatz. Furthermore, the arbitrariness of the weight function also reflects the freedom we have to choose wavefunctions representing the initial state.

5.3 Fourier series decomposition

Let us study the eigenvalue problem for a phase space distribution of the form³ $f(x, v) = \rho(x)\delta(v)$, meaning the product of a generic distribution in space with a delta function in velocity space. This choice corresponds to the case of CDM at early times, when the velocities are negligible.

³ For the sake of simplicity we restrict the analysis of this section to the one-dimensional case, but the generalization to the three-dimensional case is straightforward.

In such a case, the integral operator $g(x_-, x_+)$ becomes real and symmetric

$$g(x_-, x_+) = \rho \left(\frac{x_+ + x_-}{2} \right). \quad (68)$$

We choose the trivial weight function $w(x) = 1$, which might not be the optimal choice for a minimal number of wavefunctions, but yields a working example of the method. We now assume a periodic distribution of matter in $[0, L]$ and expand the density as a Fourier series over the interval

$$\rho(x) = \rho_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n}{L}x\right) + \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi n}{L}x\right). \quad (69)$$

The term ρ_0 can be dropped without loss of generality as it can trivially be represented in the WDF using a constant wavefunction.

The eigenvalue problem is easier to solve on the doubled interval $[0, 2L]$. The standard scalar product for real functions on this interval is simply

$$\langle \psi | \phi \rangle := \frac{1}{L} \int_0^{2L} \phi(x) \psi(x) dx, \quad (70)$$

which means that the eigenvalue problem reads

$$\frac{1}{L} \int_0^{2L} \rho \left(\frac{x+y}{2} \right) \psi(y) dy = \lambda \psi(x). \quad (71)$$

We now have to choose an orthonormal basis for the wavefunctions ψ . As we work with a periodic interval, it is natural to use harmonic functions over $[0, 2L]$. The most general case is thus

$$\psi(x) = \sum_{n=1}^{\infty} \left[\alpha_n \cos\left(\frac{\pi n}{L}x\right) + \beta_n \sin\left(\frac{\pi n}{L}x\right) \right]. \quad (72)$$

Using trigonometric identities and the orthonormality relations between the sine and cosine functions of different modes, the problem can be recast in a matrix problem for the coefficients of the Fourier series:

$$\begin{pmatrix} a_n & b_n \\ b_n & -a_n \end{pmatrix} \begin{pmatrix} \alpha_n \\ \beta_n \end{pmatrix} = \lambda \begin{pmatrix} \alpha_n \\ \beta_n \end{pmatrix}. \quad (73)$$

Therefore, the normalized eigenfunctions and eigenfunctions of the integral operator are finally given by

$$\psi_n^{\pm}(x) = \mathcal{N} \left[(a_n \pm \lambda_n^{\pm}) \cos\left(\frac{\pi n}{L}x\right) + b_n \sin\left(\frac{\pi n}{L}x\right) \right], \quad (74)$$

$$\lambda_n^{\pm} = \pm \sqrt{a_n^2 + b_n^2}, \quad (75)$$

where $\mathcal{N} = [(a_n \pm \sqrt{a_n^2 + b_n^2})^2 + b_n^2]^{-1/2}$ normalizes the eigenfunctions to unity. It can be checked explicitly that these eigenvectors satisfy the condition of orthonormality and yield the correct spectral representation

$$\rho \left(\frac{x_+ + x_-}{2} \right) = \sum_{n=1}^{\infty} [\lambda_n^+ \psi_n^+(x_+) \psi_n^+(x_-) + \lambda_n^- \psi_n^-(x_+) \psi_n^-(x_-)] \quad (76)$$

corresponding to the WDF

$$P_W(x, v) = \sum_{n=1}^{\infty} \int e^{\frac{i}{\hbar}vy} \left[\lambda_n^+ \psi_n^+ \left(x + \frac{y}{2} \right) \psi_n^+ \left(x - \frac{y}{2} \right) + \lambda_n^- \psi_n^- \left(x + \frac{y}{2} \right) \psi_n^- \left(x - \frac{y}{2} \right) \right] d^3y. \quad (77)$$

As a conclusion we have been able to solve the eigenvalue problem on the finite interval and use it to find the wavefunctions for the WDF Ansatz. This applies for a generic density profile $\rho(x)$ periodic on $[0, L]$ and a phase space distribution of the form $f(x, v) = \rho(x)\delta(v)$. The wavefunctions are harmonic functions with increasing velocity. In general, we would need an infinite number of wavefunctions to avoid smoothing the smallest scales of the power spectrum. For many applications a finite or even small number of wavefunctions may be sufficient.

In this procedure, we used the geometry of the problem to decide which orthonormal basis to use. The periodicity of the density distribution naturally led us towards the use of harmonic functions. In cases where the density is not periodic, one could use Chebyshev polynomials or any other basis whose geometry helps reduce the number of modes.

As already mentioned, the technique presented in this section holds for any power spectrum and in particular is well suited to the case of WDM without initial velocities as is usually done in numerical simulations. This truncated CDM power spectrum can easily be decomposed in a Fourier series and hence used in our framework. If the thermal velocities of the WDM particles have to be included, then another technique has to be used (see Sections 5.2 and 5.5).

5.4 Cosmological initial conditions

Observations of structure in the Universe are perfectly compatible with the simplest possible statistical description, namely a Gaussian distribution. More precisely, each Fourier mode of the density contrast $\delta(\mathbf{k})$ (not to be confused with the Dirac delta distribution) satisfies an isotropic Gaussian distribution, entirely described by the power spectrum $P(k) := \langle |\delta(\mathbf{k})|^2 \rangle$, which is a function of the modulus k only, not of the direction. From the knowledge of the power spectrum one can then generate a realization with the desired statistical properties

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \left[\sqrt{P(k)} \mathcal{N}(0, 1) \cos(\mathbf{k} \cdot \mathbf{x}) + \sqrt{P(k)} \mathcal{N}(0, 1) \sin(\mathbf{k} \cdot \mathbf{x}) \right], \quad (78)$$

where $\mathcal{N}(0, 1)$ denotes a Gaussian random number with zero mean and unit dispersion. This shows that the density contrast for cosmological initial conditions is in a form for which we know how to construct the WDF, provided that we start our simulation at times, when the Zel'dovich velocities of the particles are negligible. Compared to N -body simulations we do not need to first perform a FFT to compute $\delta(\mathbf{x})$ but can find the initial wavefunctions directly from the power spectrum. Additionally, we do not need any glassy pre-initial conditions to model the constant background.

There is, however, a little caveat when generating initial conditions for CDM. Such an initial spectrum is formally made of a Dirac distribution in v -space which means that even an infinite number of continuous wavefunction cannot reproduce exactly this singularity. This can also be explained by the quantum aspect of our formalism. Heisenberg's uncertainty principle forbids us to have at the same time an infinitely precise description of position and velocity of our wavefunction. There will be some necessary spread in velocity space proportional to the value of \hbar chosen in the simulation. The spectrum obtained will thus formally not exactly be the CDM one

but will contain some intrinsic velocities for the DM. These would vanish in the limit $\hbar \rightarrow 0$.

We would in principle require as many wavefunctions as Fourier modes are relevant in the power spectrum, which may lead to a prohibitive computational cost. Expanding the power spectrum in another basis or using a non-trivial weight $w(x)$ in the scalar product (67) may help reduce the number of wavefunctions required. On the other hand, we may turn this as an advantage as this new formalism can allow us to probe some parts of the power spectrum only without having to use the full range of \mathbf{k} .

5.5 Matrix formulation

Given that the WDF Ansatz can be thought of as spectral decomposition of an Hermitian operator, we can now analyse the solution in the discrete case, where the problem reduces to a matrix problem. Let us again restrict the analysis to one dimension. Working on a lattice (x_1, x_2, \dots, x_M) , we can think of any function $f(x)$ as a vector $(f(x_1), f(x_2), \dots, f(x_M))^T$ and of any function of two variables as a matrix. We can thus reinterpret the functional relationship

$$g(x_-, x_+) = \sum_{n=1}^N \lambda_n \psi_n^*(x_-) \psi_n(x_+) \quad (79)$$

in terms of matrices

$$\hat{G}_{ij} = \sum_{n=1}^N \lambda_n \Psi_{jn}^* \Psi_{in} = \sum_{n=1}^N \sum_{k=1}^N \Psi_{in} \lambda_n \delta_{nk} \Psi_{kj}^\dagger. \quad (80)$$

The property $g(x_+, x_-) = g^*(x_-, x_+)$ translates into the fact that $\hat{G} \in \mathbf{C}^{M \times M}$ is a Hermitian matrix $\hat{G}^\dagger = \hat{G}$ which we can diagonalize by means of a unitary transformation

$$\hat{G}_{ij} = (\Psi \cdot \hat{\Lambda} \cdot \Psi^\dagger)_{ij}, \quad (81)$$

where

$$\Psi \in \mathbf{C}^{M \times N}, \quad \hat{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N) \in \mathbf{R}^{N \times N}. \quad (82)$$

The columns of Ψ are the wavefunctions ψ_n sampled on the lattice. The property that Ψ is unitary $\Psi^\dagger \Psi = \mathbf{1}$ implies that the normalization of the wavefunctions on the lattice. This matrix formulation has the advantage, that it is straightforward to compute the spectrum of any given Hermitian matrix. The shortcomings of this approach are two-fold: first we would need as many wavefunctions as lattice points, which comes at a big computational cost, and secondly the eigenvectors have no a priori reason to be smooth enough to be used as initial conditions for our numerical scheme. Note, however, that in all the cases we tested, the eigenvalue decomposition has yield smooth enough functions.

Moreover, it has to be noted that we would need to compute the eigenvectors for a matrix containing the full three-dimensional lattice. Computing the eigenvectors of an $n \times n$ matrix is in general a problem of complexity $\mathcal{O}(n^3)$. Since the size of the matrix is related to the number of lattice points M^3 , one quickly reaches such lattice sizes making the solution of the eigenvalue problem impossible. This issue can be solved by combining this technique with the minimization procedure. One can first use an eigenvalue decomposition on a coarse grid and use this as an input of the brute-force minimization algorithm on a finer grid. A technique using multiple grids at the same time could also be used in the same way as Gauss–Seidel relaxation is done in some particle-mesh gravity solvers.

There are multiple known algorithms available to decompose a matrix in eigenvectors. We chose to use the singular value decomposition (SVD) as the publicly available implementations return the eigenvalues sorted in decreasing order. This allows us to choose only the wavefunctions whose eigenvalues are above a certain (arbitrarily chosen) level.

5.6 Discussion and remarks

For numerical simulations in a finite box with periodic boundary conditions, the spatial lattice resolution also dictates the resolution in velocity space. The size of the box is related to the lattice size in v -space since the wavevectors take discrete values $\mathbf{v} = \frac{2\pi}{L} \mathbf{n}$. The maximal wavevector is linked to the lattice spacing in real space. This illustrates the relationship between the number of wavefunctions and the spatial resolution of the simulation. If we keep all the modes, we need $\mathcal{O}(M^3)$ wavefunctions, where M is the number of lattice points in one direction. Note that this corresponds, in order of magnitude, to the number of particles in N -body simulations. So even if we keep the maximal number of wavefunctions needed to accurately represent the initial conditions, the complexity of our numerical scheme will still be comparable to the naive $\mathcal{O}(N^2)$ complexity of N -body simulations. As we will generally use much less wavefunctions, the complexity is much lower and may even trump the usual $\mathcal{O}(N \log N)$ complexity offered by tree-codes or FFT schemes to solve Poisson equation.

Another advantage of working with harmonic wavefunctions to represent the initial conditions is that we have an intuitive picture of what happens if we remove some modes. In analogy with the Fourier series, the density will not be represented exactly at every point, but the approximation becomes closer and closer as we include more and more modes. Knowing some of the properties of the system we want to model may help to get a deeper insight into which modes are really needed. The same is true when the density is expanded in another basis even if it may be more difficult to get an intuitive mental picture of the impact of high-order modes when dealing with Chebyshev polynomial say.

In many simulations, one does not necessarily need the same resolution on all scales. Instead one could work with an adaptive grid (Plewa, Linde & Weirs 2005) and have higher resolution in the scales of interest. This would allow us to reach better precisions while keeping the number of wavefunctions constant. A similar technique is used in N -body solvers such as RAMSES (Teyssier 2002) or ART (Kravtsov 1999).

In the special case of simulations of cosmic structure formation, the concept of cosmic variance could help to further reduce the number of wavefunctions required. Indeed, given that we can only observe one universe, the statistical fluctuation in large angular patches is high, as not many statistically independent patches are available in our sky. This is a well-known fact when studying the CMB radiation. This means that the statistical error is anyway large on these scales, so we do not need to work with a very high precision. Let us also recall that the freedom of choosing the weight function in the scalar product (67) of the eigenvalue problem may help to considerably reduce the number of wavefunctions. Even though this seems to be a promising route to take, we did not investigate it any further in this work.

Another area of interest could be the derivation of a scheme to generate initial wavefunctions analytically in the case of WDM (see for instance, Boyarsky, Ruchayskiy & Shaposhnikov 2009) or for any initial distribution with non-zero initial velocity spread.

6 IMPLEMENTATION AND NUMERICAL RESULTS

In the previous two sections, we showed how the cosmological V-P problem (3) can be approximated by the Schrödinger–Klein–Gordon system (49). We showed that this approximation is valid in the limit $\hbar \rightarrow 0$, $c \rightarrow \infty$, $N \rightarrow \infty$. We also demonstrated how the wavefunctions can be built and that in general they can approximate the true density distribution in the limit $N \rightarrow \infty$. For some specific cases or for smart choices of eigenfunction basis, the exact $f(x, v)$ can even be ensured with a finite or low N . But let us keep the general case in mind.

Contrary to the N -body framework, where the convergence towards the exact solution is not granted in general, we propose a method where we have a handle on the behaviour of the simulation and where we are able to easily test the dependency of the result on the parameters \hbar , c and N . This allows us to truly speak about converged results and understand the limits of our model.

Let us now present how this scheme can be discretized and implemented on a computer. We will present the implementation we used, which is probably the simplest version of what can be done.

6.1 Implementation

The simplest possible numerical scheme to solve partial differential equations is to use an explicit scheme in time. An implicit scheme would be more precise but would require more computing time and memory, the latter quantity being, as we will show, a rather scarce resource. This explains the choice of an explicit scheme, even if this imposes the use of a Courant-like condition for our time steps. For the same reasons, a scheme accurate up to order $(\Delta\tau)^2$ in time has been chosen. As going to a precision of order $(\Delta\tau)^4$ would require almost twice as much memory, this choice can reasonably not be made. Using a symplectic integrator may, however, be useful in future studies as they do not cost more in terms of memory but conserve the energy of Hamiltonian systems exactly.

Regarding the spatial derivatives, there are no constraints coming from the memory requirements. One could in principle go to an arbitrary level of accuracy. But as the time derivatives only have a limited precision, it is not worth going to a precision higher than $(\Delta x)^4$, using the usual five-point stencil.

With these two points being set, the system of equations (49) can be written on a lattice as follows:

$$\begin{aligned} \psi_n(x, \tau + \Delta\tau) &= \psi_n(x, \tau - \Delta\tau) + i \frac{\hbar \Delta\tau}{a(\tau)} \nabla_{\text{dis}}^2 \psi_n(x, \tau) \\ &\quad - i \frac{2\Delta\tau}{\hbar} U(x, \tau) \psi_n(x, \tau) \\ U(x, \tau + \Delta\tau) &= 2U(x, \tau) - U(x, \tau - \Delta\tau) \\ &\quad + c^2 \Delta\tau^2 \nabla_{\text{dis}}^2 U(x, \tau) \\ &\quad - \frac{4\pi G c^2 \Delta\tau^2}{a(\tau)} \left(\sum_n^N \lambda_n |\psi_n(x, \tau)|^2 - \bar{\rho} \right), \end{aligned}$$

where the discretized divergence operator is given by

$$\begin{aligned} \nabla_{\text{dis}}^2 f(x) &= \frac{1}{12\Delta x^2} [-f(x + 2\Delta x) + 16f(x + \Delta x) \\ &\quad - 30f(x) + 16f(x - \Delta x) - f(x - 2\Delta x)]. \end{aligned}$$

In the non-cosmological case, the factors $a(\tau)$ can be dropped and one can use time t instead of conformal time τ . One can show that

this scheme is unitary and conserves the norm of each wavefunction. Since the iterative solution contains the fields at neighbouring lattice sites, care has to be taken that the boundary conditions are implemented correctly. This is most easily done by augmenting the arrays containing the values of the fields on the lattice by so-called ghost points to store the periodic boundary conditions.

The last important point regarding the numerics is the choice of c and \hbar . It is clear that the Klein–Gordon equation reduces to the Poisson equation in the limit $c \rightarrow \infty$ and that the higher order terms of (39) vanish in the limit $\hbar \rightarrow 0$. But numerical stability imposes more conditions on these values. An explicit scheme can only converge if there is no information propagating a distance of one cell during one time step. The scalar field propagates at speed of c , which gives us the following condition:

$$\frac{\Delta x}{\Delta\tau} > c, \quad (83)$$

which is the usual *Courant condition*. In practice, the right-hand side is multiplied by a constant (10–10²) in order to avoid any instability and to remain far from the actual condition. This condition gives a clear relation between those three quantities and shows that one cannot arbitrarily improve the spatial discretization without changing the time step size. It is not surprising to have to introduce such a condition. Indeed, if we were to truly use a value of $c = \infty$ in our simulations, we would have to use smaller and smaller time steps for a fixed grid spacing. At some point, solving the Poisson equation would become algorithmically cheaper. The Courant condition is thus the price to pay to avoid solving the usual Poisson $\mathcal{O}(M \log M)$ problem.

The evolution of the Schrödinger equation also imposes conditions on the time and space slicing. It can be shown that the following relation

$$\frac{\Delta x^2}{\Delta\tau} > \hbar \quad (84)$$

must hold, encouraging us, once again, to choose \hbar as small as possible. At this stage, no lower bound has been analytically derived for \hbar . The full dependence on \hbar of the simulation results is still an open question left for further investigation of this framework.

6.2 Complexity and memory requirements

Having presented the algorithm of the time evolution, let us estimate its computational complexity and memory requirements. Consider a three-dimensional spatial grid made of M^3 lattice points. Let N_ψ be the number of wavefunctions we evolve. Adding the spatial components of the scalar field, $N_f = N_\psi + 1$ is the total number of fields we evolve in time. At each time step, we need to compute each of the fields at every lattice point, making the algorithm of complexity

$$\mathcal{O}(M^3 \cdot N_f). \quad (85)$$

This has to be compared with N -body simulations, which have a naive complexity of $\mathcal{O}(N^2)$, that can be reduced to $\mathcal{O}(N \log N)$ using optimized algorithms. The more particles are tracked, the better becomes the spatial resolution. Roughly, for a total of N particles, $\Delta x_{\text{resol}} \sim L_{\text{box}}/N^{1/3}$. In our case, the spatial resolution is defined by the lattice spacing $\Delta x_{\text{resol}} \sim L_{\text{box}} M^{-3}$. Thus, for comparable spatial resolution, we would need $M \sim N^{1/3}$. From this, we conclude that the complexity of our algorithm scales as $\mathcal{O}(M \cdot N_f)$. In the ideal situation, where we only need a few wavefunctions, $N \sim \mathcal{O}(1)$, our new framework provides an $\mathcal{O}(M)$ algorithm to

study structure formation. It seems that in the worst case we would need as many wavefunctions as there are Fourier modes on the lattice, $N_\psi \sim \mathcal{O}(M^3)$ implying a complexity $\mathcal{O}(M^2)$, which is the same as the naive force summation in N -body simulations.

These estimates illustrate that our algorithm can indeed compete with the complexity of N -body simulations. It also shows how crucial it is to reduce the number of wavefunctions as much as possible.

Let us next have a look at the memory requirements of our approach. Given that our time evolution relies on a two-level explicit scheme, we need to keep the field configurations at two time steps in memory. For N_ψ complex wavefunctions and one real scalar field components on the whole lattice, we need $2 \cdot M^3(2N_\psi + 1)$ variables. Assuming that each is stored as a double of 8 bytes, we can estimate the minimal memory needed by our numerical simulation to be

$$\geq 2 \cdot M^3(2N_\psi + 1) \cdot 8 \text{ bytes.} \quad (86)$$

Let us look once more at the worst case scenario $N_\psi \sim \mathcal{O}(M^3) \sim \mathcal{O}(N_\psi)$. Hence, the memory required now raises to

$$\geq 32 \cdot N_\psi^2 \text{ bytes.} \quad (87)$$

This has to be compared with N -body simulations, which have to store at least the position and velocity of each particle at every time step leading to a memory consumption of

$$\geq 2 \cdot 6 \cdot N \cdot 8 \text{ bytes.} \quad (88)$$

As an example we may give the Millennium simulation (Springel et al. 2005), which needed about 400 GB to store the information of their $2160^3 \simeq 10^{10}$ particles, in agreement with the above estimate. We have to conclude that our approach can be strongly constrained by its memory requirements. The gain in computational complexity seems to have come at a considerable cost in memory. If we consider the 1 TB of memory available to the Millennium simulation, we could only have $\sim 57^3$ lattice points! However, if we were to use as many wavefunctions as spatial lattice points, we could as well directly simulate the V-P system without introducing any approximation. The whole point of the framework we introduced is to simulate a realistic probability distribution with a low number of wavefunctions, in which case the memory requirements are not prohibitive any more and scale with N as in the N -body case. We also mentioned the idea of using an adaptive mesh to improve the (spatial) resolution without having to increase the number of wavefunctions.

We now turn to two cases we simulated and show that this new framework is able to reproduce the known solutions. We also show how the solution depends on the parameters c , \hbar , N and Δx .

6.3 Spherical collapse of a DM sphere

There are few known non-trivial analytical solutions to the V-P system (3) even in the static Universe ($a(\tau) = 1$) case. The collapse of a uniform sphere is among these and is of particular interest for cosmological applications. A comprehensive treatment of the case, known as *Tolman solution* (Tolman 1934), can, for instance, be found in the textbook (Weinberg 1972). A uniform sphere of initial density ρ_0 and radius R_0 is collapsing under its own gravitational potential. Gauss' law for gravity states that the evolution of a sphere is not influenced by the matter lying outside itself. This means that the density inside the sphere will remain constant with the radius at every time t . In other words, all matter will reach the centre at the same time which will lead to an infinite density. At this stage, the Newtonian description becomes invalid and one would

have to use general relativity (GR) in order to take into account all the effects. In the framework of Newtonian gravity, the matter will simply cross the centre and oscillates around the centre. Due to the discretization, the simulated central density cannot become infinite and these oscillations cannot be reproduced exactly. The same shortcomings are present in N -body codes.

The evolution of the radius R with time is a quantity which can be easily tracked. In parametric form, the Tolman solution reads ($0 \leq \beta \leq \pi$):

$$t = \frac{\beta + \sin \beta}{2\sqrt{\frac{8\pi G}{3}\rho_0}}, \quad (89)$$

$$R = \frac{1}{2}(1 + \cos \beta). \quad (90)$$

The density inside the sphere will evolve following the relation:

$$\rho(r, t) = \frac{\rho_0 R_0^3}{R^3(t)}. \quad (91)$$

For simplicity in what follows, we set $R_0 = 1$, $G = 1$ and $\rho_0 = \pi$. The final collapse time (in arbitrary units) is then reached when $t_c \approx 0.306$. We will work on the periodic interval $[-5, 5]$ which should be big enough to avoid any unwanted effects from the boundaries.

This problem possesses an obvious spherical symmetry and in order to be able to explore a wide resolution range it is interesting to rederive the whole framework presented in Sections 4 and 5 using this assumption. A careful derivation can be found in appendix A and the end result is that the V-P system with spherical symmetry can be recast in the one-dimensional Schrödinger–Klein–Gordon system

$$i\hbar \frac{\partial \psi_n}{\partial \tau} = \frac{-\hbar^2}{2a(\tau)} \frac{\partial^2 \psi_n}{\partial r^2} + \frac{V}{r} \psi_n, \\ -\frac{1}{c^2} \frac{\partial^2 V}{\partial \tau^2} + \frac{\partial^2 V}{\partial r^2} = 4\pi G r \left(\frac{2\pi}{r^2} \sum_n \lambda_n |\psi_n|^2 - \frac{4\pi^2 \Xi}{V_{\text{tot}}} \right),$$

where the potential $V = Ur$ and Ξ is the normalization of the wavefunctions (see equation A12). The main difference with the framework presented earlier is the explicit dependency on the position coordinate r . The algorithms developed to find the wavefunctions corresponding to a given distribution function are identical.

To generate the initial set of wavefunctions and eigenvalues we chose to use the matrix formulation (Section 5.5). The initial density profile being discontinuous, it is obvious that it cannot be recovered exactly with a finite set of continuous functions. There will be some noticeable differences between the exact density profile and its approximation appearing at the discontinuity points, that is at the edge of the sphere. It is thus better to use an approximately correct but continuous density profile. In the case at hand, we used the following initial setup:

$$\rho(r, t = 0) = \frac{\pi}{2} \tanh(\xi(r + 1)) - \frac{\pi}{2} \tanh(\xi(r - 1)), \quad (92)$$

with $\xi = 20$. The value of ξ is somewhat arbitrary and has been chosen in order to be as close as possible to the perfect sphere (i.e. high ξ) and avoid any Gibbs oscillation at the edge of the sphere (i.e. low ξ). The results presented here are not really dependent on ξ . This parameter has just been introduced for convenience and to avoid having to analyse the effects of these unwanted and unrealistic oscillations. In fact, even a value of $\xi = \infty$ yields comparable results to what is shown below once the Gibbs oscillations have been smoothed out manually from the output.

Once discretized on a lattice, the eigenvalue decomposition is straightforward to obtain, for instance using the SVD function implemented in the usual scientific software packages. Recall that there is no guarantee that the obtained functions will be periodic on the interval of interest or even that these functions will be smooth. It is a pure matrix operation without any relation between the matrix elements representing the wavefunctions. The interval $[5, 5]$ has been uniformly discretized regularly in 5000 line elements in order to get a high enough spatial accuracy. This means that we want to perform the SVD decomposition of a 5000×5000 matrix and that we can use up to $N = 5000$ wavefunctions in the simulation. The matrix reads

$$\hat{G}_{ij} = \rho \left(\frac{r_i + r_j}{2} \right), \quad (93)$$

where the r_i s are the uniformly distributed lattice points. This matrix is by construction symmetric and positive definite, meaning that its eigenvalues will be positive or null. Most of the SVD routines in scientific packages sort the eigenvalues λ_n according to their magnitude which allows us to classify the most important contributions and discard the negligible terms in equation (27) if one does not want to use all the N functions. The first four wavefunctions are shown in Fig. 1.

The wavefunctions obtained through this procedure are smooth (at the lattice level at least) and real but are not periodic nor anti-periodic, which leads to spurious diffusion at the boundaries of the box. For this reason, we decided to multiply them by a square-box like compact function going to zero close at the box boundaries. The first four wavefunctions before and after applying this window filter are also shown in Fig. 1. This procedure does not modify the distribution function obtained through the WDF. This reflects the fact that there is infinitely many ways to decompose the same $f(r, v)$ in wavefunctions. Notice that this procedure of adding a window function can only be done if the density vanishes at the boundaries.

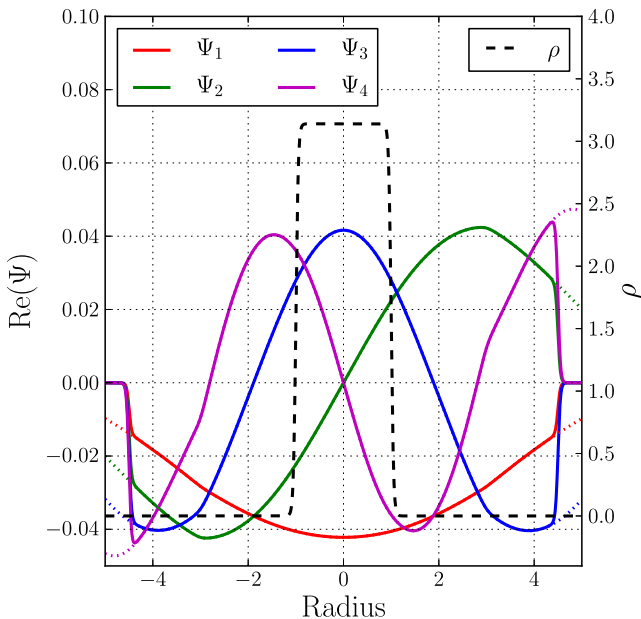


Figure 1. The first four wavefunctions contributing to the WDF of the approximate uniform sphere before (dotted lines) and after (superimposed solid lines) having applied the smooth window function to make them vanish at the boundaries of the box. These functions are different from zero almost everywhere but their combination in a WDF corresponds to the density profile (dashed black line, equation 92), which is zero on most of the interval.

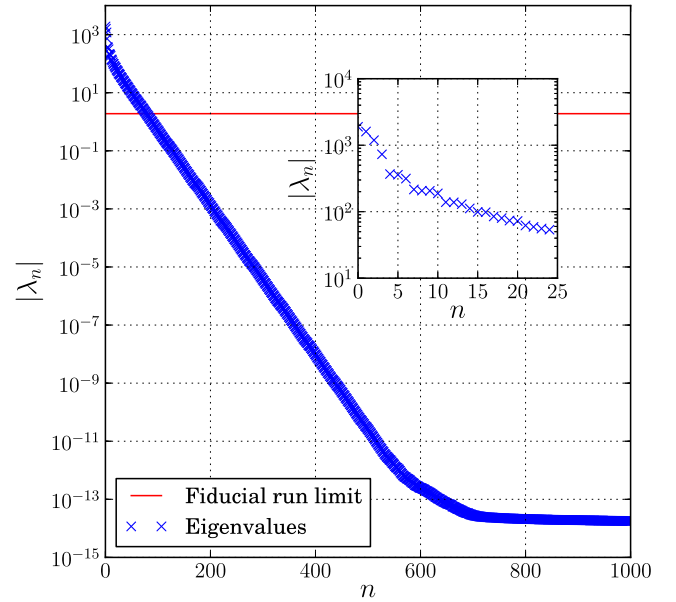


Figure 2. The first 1000 eigenvalues λ_n corresponding to the SVD decomposition of the spherical collapse problem. The values decrease rapidly and become negligible (when compared to the first one) for $n > 100$. They even reach a minimum close to the machine epsilon for $n > 700$. Our fiducial run uses all the wavefunctions up to $n = 79$ which corresponds to $\frac{\lambda_n}{\lambda_0} > 10^{-3}$. This limit is shown as the red solid line in the figure. The small panel presents a zoomed-in region of the eigenvalues with $n < 25$. The decrease on this small subset is already of more than an order of magnitude.

Apart from the wavefunction, the eigenvalue associated with each mode also enters the WDF (equation 27). These are obtained at the time than the discretized wavefunctions and their values are represented in Fig. 2. The actual normalization of the eigenvalues does not really matter as any common factor can be absorbed as normalization in front of the WDF. But the ratio of the values plays a role. All the different wavefunctions (modes) entering the decomposition of $f(r, v)$ may not play an important role exactly as in the case of a Fourier series decomposition where some of the modes can safely be neglected. As can be seen in Fig. 2, the values of the various λ_n decrease rapidly and for $n > 100$, they represent less than 10^{-3} of the most important mode. As the eigenvalues are constants of motions, we can hope that neglecting modes with a high n (and hence a small λ_n) will not affect the simulation too much. In fact, unless the magnitude of the wavefunction corresponding to one of the neglected mode grows significantly over the course of the simulation, this mode should remain small at all time and can thus be safely ignored.

In our main run, we used all eigenfunctions Ψ_n whose eigenvalue fulfils $\lambda_n > 10^{-3}\lambda_0$, which left us with only $N = 79$ functions to evolve. The other numerical parameters we chose in our fiducial run are $c = 10$ and $\hbar = 0.005$. We do not expect \hbar to have a big impact on the results in this case as the potential is a combination of a second and third order polynomial for which the higher order corrections in the Wigner equation (39) should be small.

Fig. 3 shows four density profiles at different time steps in the simulation together with the analytical solution (equation 91). Until $t \approx 0.2$, the behaviour of the density profiles remains close to the exact solution apart from the very edges of the sphere that are slightly smoothed. The centre of the density profile is almost flat as expected and has almost the correct value. When coming closer to the collapse time $t_c \approx 0.306$, the profiles starts to deviate more

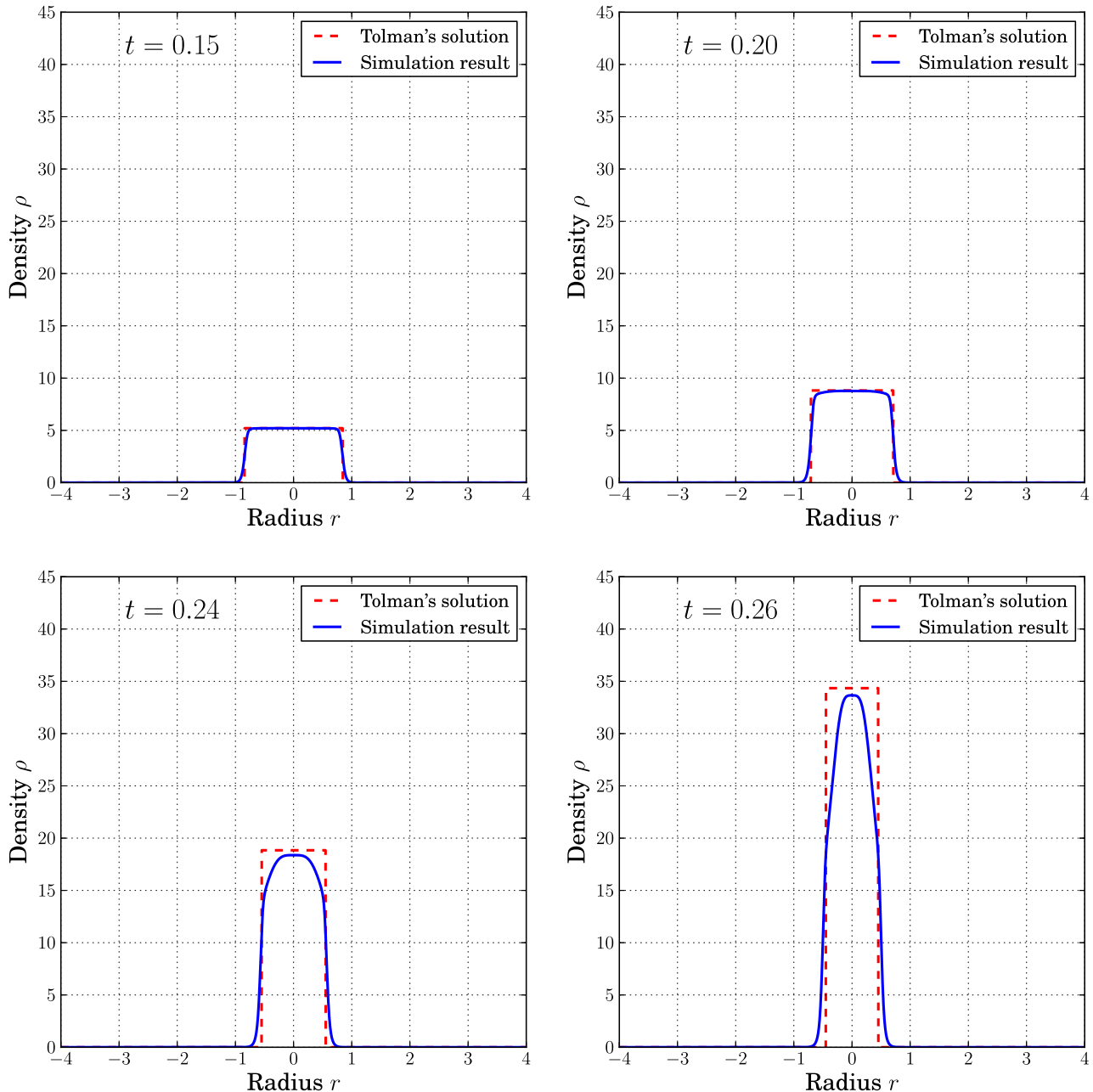


Figure 3. Four different density profiles (blue solid line) of our fiducial run at $t = 0.15, 0.20, 0.24$ and 0.26 together with the analytical solution (red dashed line) at the same time steps. The simulation follows almost perfectly the analytical solution until $t \approx 0.20$ and then starts to deviate. For obvious reasons, the situation gets worse at later times and even the centre of the sphere does not follow the exact solution any more. After the collapse time $t_c \approx 0.306$ the behaviour becomes clearly non-physical due to the unresolved infinities. The very centre of the sphere still follows Tolman's solution closely on the last two panels but the sharp features at the edge of the sphere get increasingly more difficult to represent. This suggests that increasing the spatial resolution of the lattice may help getting better derivative estimate and hence improve the quality of the result.

and more from the expected profile. This can be seen on the last two panels of Fig. 3 where the density inside the sphere is clearly different from a square box function. The very centre of the sphere still remains close to the analytical solution but the edges are not sharp any more and are smoothed over many lattice elements. This strongly suggests that the estimation of the derivatives of both the potential and the wavefunctions are getting poor or that the number of wavefunctions used in the run is not high enough. Our scheme uses a fourth order accurate derivative stencil but this does not necessarily help recovering sharp features such as the one present

at the edge of the sphere. Increasing N and reducing Δr may help recover the right density profile everywhere in the sphere.

The results in Fig. 3 have been obtained using $N = 79$ wavefunctions corresponding to all eigenvalues $\lambda_n > 10^{-3}\lambda_0$. This should be sufficient as the eigenvalues are constants of motion and we do not expect any of the neglected wavefunctions to grow by a huge factor over the course of the simulation. In order to assess this, we run the same simulation with $N = 155$, corresponding to all wavefunctions whose eigenvalues $\lambda_n > 10^{-5}\lambda_0$. Notice here that decreasing the minimal eigenvalue entering the WDF by two orders of magnitude

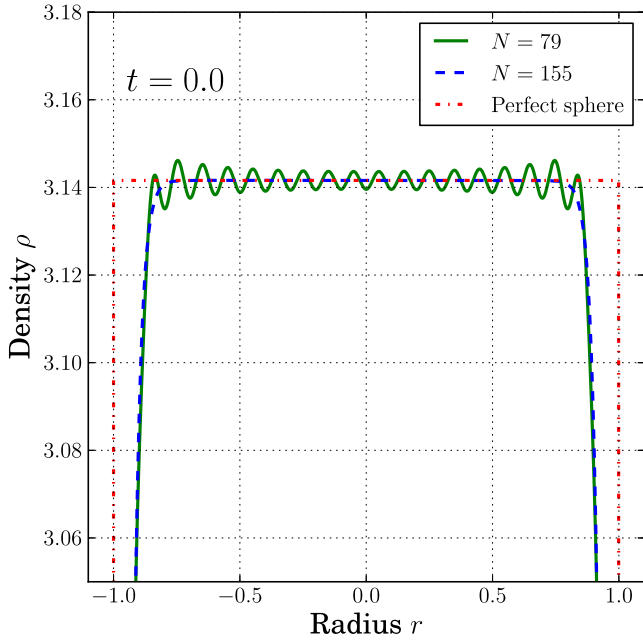


Figure 4. Comparison of the density profiles at the initial time for $N = 79$ (green solid line) and $N = 155$ (blue dashed line) wavefunctions. The figure zooms in the central regions where the difference can be spotted. The $N = 79$ line presents a lot of oscillations that are suppressed if more wavefunctions are used. The $N = 155$ line almost perfectly matches the density profile given by equation (92). Notice, however, that this differs from the perfect sphere profile (red dash-dotted line), which cannot be represented by a finite set of continuous functions.

only increases N by a factor of 2. We are thus far from the worst case scenario (see Section 6.2) where the same number ($N = 5000$) of wavefunctions than lattice points have to be used.

Fig. 4 shows a comparison at $t = 0$ of those two initial setups. The figure only shows a zoomed-in view focused on the sphere itself as the differences are less visible in the outer regions of the simulation domain. As can be seen, the $N = 155$ initial setup (dashed blue line) is a much better representation of the smoothed density profile (equation 92). At this resolution, the two are indistinguishable. The $N = 79$ initial conditions (green solid line) present some oscillations inside the sphere that are very similar to the Gibbs phenomenon that appears when computing the Fourier series of the square box function. Using a smoothed density profile and decomposing in eigenvalues using the matrix formulation thus yields a result which is very similar to generating the initial conditions (ICs) through the Fourier decomposition (Section 5.3). This could have been anticipated by looking at the wavefunctions (Fig. 1), where the different ψ s resemble sines and cosines functions at least qualitatively. As can be seen, the relative error introduced by using only $N = 79$ wavefunctions is of the order 10^{-3} , whereas the error computed when using $N = 155$ is smaller than 10^{-6} , showing once again that increasing the number of eigenfunctions used by a factor of 2 increases the simulation by more than two orders of magnitude. However, it should be noticed that using another basis or weighting function for the eigenvalue decomposition (67) may yield another N with the same or different accuracy. Comparing the number of wavefunctions only makes sense when using a similar decomposition technique. Let us also mention that we also tried using harmonic functions and Chebyshev polynomials for this test case and obtained similar results.

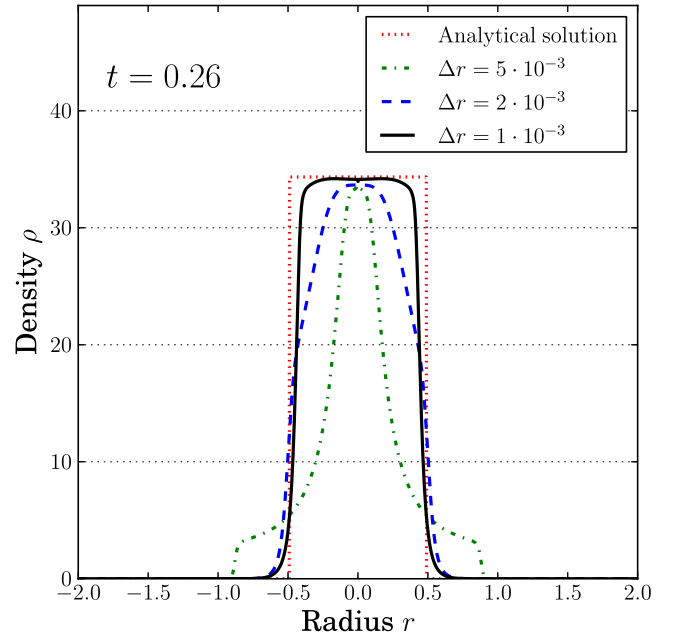


Figure 5. The output at $t = 0.26$ for different lattice resolution using $N = 79$ wavefunctions. The green dash-dotted line corresponds to a low-resolution run with $\Delta r = 5 \cdot 10^{-3}$, the blue dashed line corresponds to our fiducial run at $\Delta r = 2 \cdot 10^{-3}$ and the black solid line is the output of a high-resolution run using $\Delta r = 10^{-3}$. The quality of the output is clearly improved by using a higher resolution lattice. This can be directly related to the problem of estimating sharp derivatives on a grid, where the only solution is to increase the resolution.

At later times, the simulation snapshots are identical to the ones presented earlier in Fig. 3. The relative difference between the two runs is of the order of 10^{-3} as in the initial conditions. This implies that the difference between our simulation results and the analytical solution cannot be reduced by using more and more wavefunctions. The additional modes that have been discarded when using only 79 eigenfunctions do not contribute significantly to the final results. This could have been expected as their weightings (λ_n) are very small compared to the main modes. We can thus gain confidence in the way we generate ICs, discarding higher order modes may not be an issue and we may be able to run our algorithm in a near linear regime even when a violent collapse of matter is studied.

In conclusion, increasing N does make the initial conditions and the simulation outputs converge towards a solution at a high rate. However, the discrepancy between the solution and the simulation does apparently not come from the wrong choice of the parameter N . Let us now explore the dependency on the grid resolution.

In Fig. 5, we show the results of three runs at different grid resolutions leaving the number of wavefunctions and all the other parameters fixed. The blue dashed line corresponds to the fiducial run ($\Delta r = 2 \cdot 10^{-3}$), the green dotted line to a lower resolution run using $\Delta r = 5 \cdot 10^{-3}$ and the black solid line corresponds to the high-resolution run with $\Delta r = 10^{-3}$. As can be seen, increasing the resolution has a huge impact on the quality of the result. As anticipated, the sharp features can only be resolved correctly when enough grid points are used. Notice that the high-resolution run almost matches exactly a rescaled version of the initial density profile (equation 92), but does break down at later times in the same way as the fiducial run did between $t = 0.20$ and 0.26 (Fig. 3). Increasing the resolution is thus important to be able to retrieve all features of this somewhat artificial test case. This test case presents

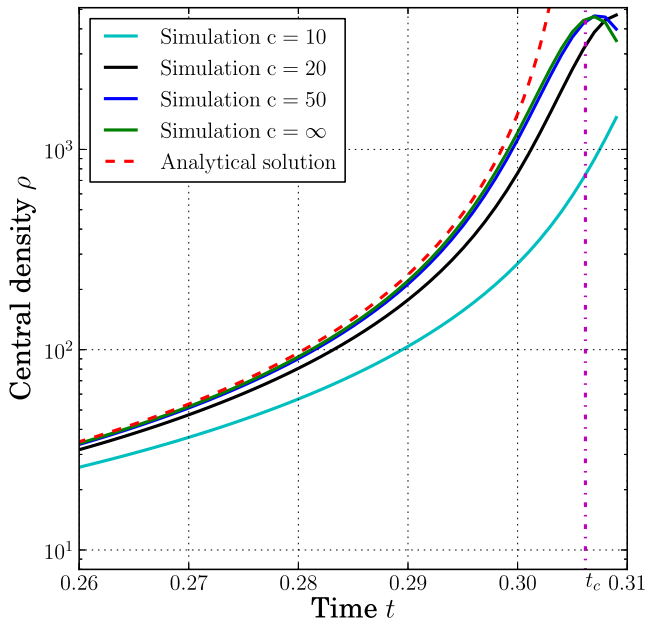


Figure 6. Evolution of the density at the centre of the sphere ($r = 0$) for different values of the numerical speed of light c . The red dashed line corresponds to the analytical solution (91), the vertical dash-dotted line represents the final collapse time and the different solid lines correspond to the different values of c . The higher the value of c , the closer the line lies to the exact solution. The line with $c = \infty$ has been obtained by solving Poisson’s equation on the grid instead of evolving gravity using Klein–Gordon’s equation. The quality of the simulation outcome clearly depends on the value of c but if the value is high enough (compared to the velocity of the matter), the difference with the $c = \infty$ becomes very small. Once the density peak has been reached, the value of $\rho(r = 0)$ decreases as is expected after the different matter shells have crossed. The simulations have not been carried on much beyond this point as the departure from the analytical solution is already significant. Moreover, the peak cannot be represented accurately by any numerical mean and any subsequent event would be erroneous.

a strong density gradient at the edge of the sphere which does not spread over many cells. This makes it difficult to resolve for a grid code but in a cosmological simulation such sharp gradients should not arise as the density profiles usually follow power laws and do not have infinite gradients. As already mentioned, using an adaptive mesh would help in such a case as more resolution elements could be used at the edge of the sphere without having to slow down the simulation due to an unnecessary oversampling of the steady regions.

This demonstrates that our framework converges towards the analytical solution once the spatial resolution is high enough and once the number of wavefunctions has been carefully chosen to represent the distribution function of interest.

This new framework should converge towards the solution in the limit $N \rightarrow \infty$, $\Delta x \rightarrow 0$, $c \rightarrow \infty$ and $\hbar \rightarrow 0$, the last two being, despite their physical origin, only numerical parameters. Fig. 6 presents the evolution of the density at the centre of the sphere for our fiducial run and for higher values of c . The simulation with $c = \infty$ has been obtained by solving Poisson’s equation on the grid at every time step instead of using Klein–Gordon’s equation. Increasing c improves the quality of the result and even relatively small values ($c = 50$) of this parameter lead to a behaviour close to the limiting case. Poisson’s equation can thus safely be replaced by Klein–Gordon’s equation. The maximal speed reached by matter

shells in our fiducial run is $p \approx 10$ before the very end of the collapse, which can anyway not be studied by a simulation. Using a value of $c = 10$ is thus intuitively too low and this plot confirms this. The speed of gravity must be at least a few times bigger than the matter velocity.

Once the peak has been reached, the different matter shells should cross the centre and the density at $r = 0$ has to decrease. The start of this behaviour can also be seen in Fig. 6. The main issue with this analysis is that is happening after the moment where the density at the centre becomes infinite and hence not representable on a computer. In practice, all the wavefunctions should become infinite at this precise point and zero elsewhere. This is obviously impossible on a lattice and does anyway lead to inaccurate derivatives. To get closer and closer to the singularity requires a finer and finer mesh. The smaller the mesh size, the better the shell crossing can be followed.

Notice, however, that this is an issue present in this ideal sphere case only. In a realistic scenario, where the matter has a non-zero radial velocity and in an expanding background, the usual NFW profiles (Navarro, Frenk & White 1996) should be recovered without singularity problems. This would, however, require a truly three-dimensional simulation and not just a spherically symmetric one-dimensional setup.

Increasing c has a big impact on the simulation run time as the time step size varies as c^{-1} , making the total simulation wall clock time proportional to c . An option that has not been explored here is to change the value of c to be always a (small) multiple of the maximal matter speed in the simulation. This would allow us to choose bigger time steps in the early stage of the simulation when all the matter moves slowly. It would also avoid making an initial guess for the value of c without knowing how fast the matter will move during the run.

As discussed earlier, the dependency on \hbar is difficult to test in this case as the analytical potential only presents first order corrections in the Wigner equation. We did run some simulations with various values of this parameter without noticing important differences in the behaviour of the matter distribution. Understanding the exact dependency on \hbar of the framework is left to a future work.

This simple spherical collapse test showed that we were able to reproduce the analytical solution in the limit $N \rightarrow \infty$, $\Delta x \rightarrow 0$ and $c \rightarrow \infty$ as expected. We investigated the different deviations from the exact solution and could explain them through our choices of numerical parameters. We also discussed how the implementation could be improved by using a mesh refinement and adaptive c values. The results obtained so far show that this new framework can reproduce known solutions and give us confidence to use it on more involved cases.

6.4 Going beyond the first collapse

With the previous test case, we showed how our framework was able to reproduce the collapse of a matter distribution in the linear regime and studied the dependency on the model parameters. However, in most cases of interest, the systems considered in simulations are way past the linear regime. They also present multiple matter streams, i.e. at a given position x , there are multiple velocities v and the distribution function is ‘wound up’. It is hence important to explore whether this behaviour can be recovered by our framework. Note that tracking precisely these multiple matter streams is extremely difficult in the case of N -body simulations unless advanced phase space tessellation techniques are used (Abel et al. 2012; Shandarin et al. 2012).

Without loss of generality, we can normalize it such that it is equal to one at the start of the simulation $a(\tau_{\text{ini}}) = 1$, implying

$$H_{\text{ini}}^2 = \frac{8\pi G}{3} \bar{\rho}_{\text{com}} a(\tau_{\text{ini}})^{-3} = \frac{8\pi G}{3} |\psi_0|^2. \quad (98)$$

The above relation fixes the value of this wavefunction in terms of the initial Hubble parameter, which can be computed by rescaling today's value H_0 to the redshift corresponding to the beginning of our simulation

$$H_{\text{ini}} = H_0(1 + z_{\text{ini}})^{3/2}. \quad (99)$$

We ran our simulations for the choice $z_{\text{ini}} = 1000$ and using today's Hubble parameter $H_0 \simeq 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The initial conditions with a density contrast of $\delta_{\text{ini}} = 10^{-6}$ where evolved up to a redshift of $z_{\text{fin}} = 200$. We use a normalized time line such that z_{ini} corresponds to $\tau = 0$ and z_{fin} corresponds to $\tau = 1$ using $3 \cdot 10^4$ time steps. The same initial perturbations were evolved in a matter-dominated, expanding universe and in a static universe without expansion.

The parameter c has been chosen in accordance with the results of the previous test by making it bigger than the speed of the matter in the simulation and small enough to avoid drastically pulling down the time step. In what follows, $c = 10$. The parameter \hbar has been, once again, chosen small enough for the quantum corrections to be negligible. More specifically, this means that the first quantum correction in the Wigner equation (39) has to be small compared to the contribution to the classical Vlasov equation:

$$\frac{\partial V}{\partial x} \frac{\partial P_W}{\partial v} > \frac{1}{24} \hbar^2 \frac{\partial^3 V}{\partial x^3} \frac{\partial^3 P_W}{\partial v^3}. \quad (100)$$

We verified that this is indeed the case in our simulations when using $\hbar = 0.005$. We could, in principle, also verify that the higher order corrections are also suppressed but computing the fifth derivative of the potential will lead to a very noisy estimate and may not lead to useful results.

Fig. 8 shows the time evolution of the density. The initial amplitude of the harmonic density increases with time, without distortion of the shape, as expected from the linear regime of structure formation. The growth of structure seems thus to be well reproduced by our framework even with such a low number of lattice points and wavefunctions. The simulation could, in principle, be carried on to a much lower redshift than $z = 200$ but at some point, the spatial resolution issues highlighted in the previous test would appear here as well. Recall that we have only 30 grid points in our 60 Mpc box. As soon as the variation of the density becomes important on a scale of the order of a few Mpc, the discretized derivatives will cease to approximate the analytical ones and our formalism will break down as would any uniform grid code with the same resolution. We, hence, decided to restrict ourselves to the regime where our density field and the wavefunctions are well behaved in order to make a useful analysis of the results.

To analyse the growth of the perturbation in more detail, we performed a Fourier transform on the density contrast to obtain $|\delta_k|^2$. In this way, we could also check that no other Fourier modes than the one initially present were excited during the simulation. This is a cross-check for the linearity of the evolution of the small density perturbation. The Fig. 9 compares the growth $|\delta_k(\tau)|^2 / |\delta_k(\tau_{\text{ini}})|^2$ for our mode in the expanding and non-expanding universes. Clearly, the growth of the perturbation is suppressed in presence of expansion.

These results clearly show that our framework is able to follow the growth of a single-mode density perturbation in an expanding background. The main features are recovered even when a low

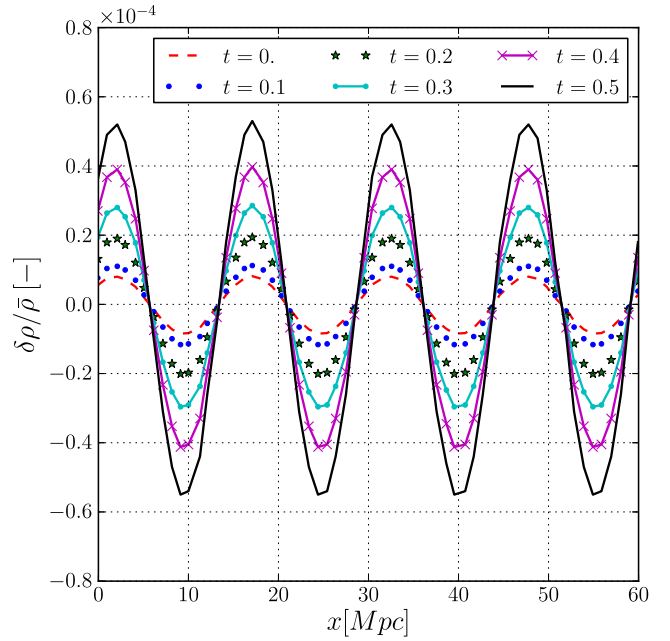


Figure 8. Time evolution of the density field in an expanding universe. The different lines correspond to various time steps in normalized units. The values are taken along one line parallel to the x -axis in the box but all lines yield the same results. The initial amplitude of the harmonic density increases with time, without distortion of the shape, as expected from structure formation in the linear regime.

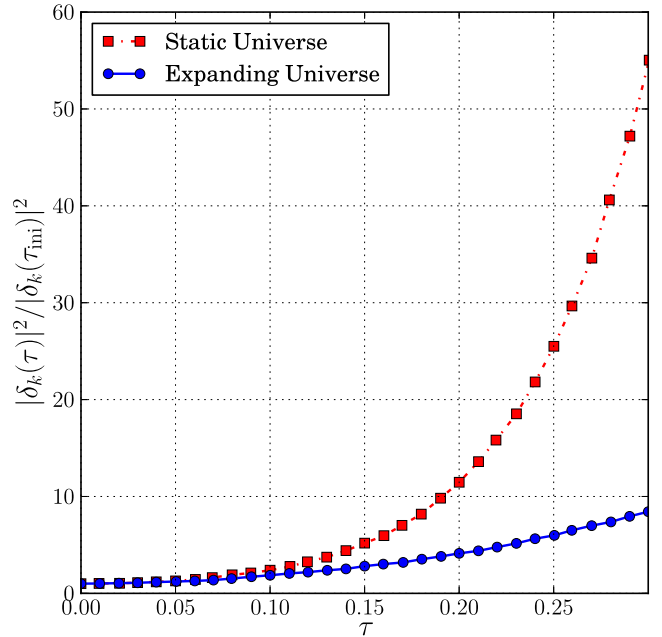


Figure 9. Comparison of the growth of the perturbation $|\delta_k(\tau)|^2 / |\delta_k(\tau_{\text{ini}})|^2$ in a non-expanding universe (red squares, upper line) and in a matter-dominated, expanding universe (blue circles, lower line) as a function of the conformal time τ . As expected, the growth is clearly suppressed in the presence of expansion.

number of lattice points and wavefunctions are used. By taking advantage of the ease of decomposition in orthonormal Fourier modes of the cosmological power spectrum (Section 5.4) more complex cases can be studied by superposing the different modes.

The results obtained here give us confidence about the behaviour of the framework in the non-linear regime of cosmic growth. The main features of Λ CDM can probably be recovered in a higher resolution run with more wavefunctions and a longer run time.

As in the previous test case, one could track the matter distribution into the non-linear regime and track the appearance of multiple matter streams. This is of course of crucial importance for realistic simulations of structure growth in the Universe. It is, however, obvious that the addition of the scalefactor $a(\tau)$ in the simulation will not alter the behaviour seen in Section 6.4 and we are confident that multiple streams would also appear and be correctly tracked by the evolution of the wavefunctions. A more detailed study of the framework in the context of cold or WDM cosmologies is left for future work.

7 CONCLUSION

We introduced a new alternative framework for simulation of structure formation which is not based on the usual discretization of the density field in a set of particles. We made use of the WDF to recast the distribution function in a set of wavefunctions. We could thus replace the six-dimensional Vlasov equation by a set of Schrödinger equations acting on the wavefunctions. The Poisson equation for gravity has been transformed into a Klein–Gordon equation making the system of equations completely local. We demonstrated how this system of equation could be derived from a Lagrangian and how the total energy and mass are conserved by the equations of motion.

We presented different methods to generate the initial conditions depending on the distribution function of interest and described how a cosmological power spectrum can be discretized in a low number of wavefunctions. The framework has then be tested on two simple models to assess its validity and the dependency of the outcome on the numerical parameters has been sketched. The results obtained thus far show that this framework is viable and may become a possible alternative to the N -body method.

The important new features introduced in this framework are the possibility to simulate a generic distribution function and not only CDM. Although finding an easy and generic way to generate initial conditions for warm or HDM remains an open question, there are no intrinsic limitations in the framework that could prevent such simulations. It also provides an alternative to N -body codes and could thus help assess the validity of simulations. Our technique can be shown to converge towards the solution in the limit $c \rightarrow \infty$, $\hbar \rightarrow 0$ and $N \rightarrow \infty$ making the formal convergence studies possible. The computational complexity of the algorithm grows as $\mathcal{O}(N \cdot M)$, where M is the number of lattice points. This demonstrates the importance of finding the appropriate decomposition of the distribution function in wavefunctions. The complexity can hence be anything between linear and quadratic in the number of points. The case of structure formation may be close to the ideal case thanks to the possibility to discretize the power spectrum in a low number of modes.

This scheme is especially aimed at tackling the fundamental challenges that the N -body method faces when dealing with non-CDM cosmologies. This includes simulation of a WDM Universe but also neutrino components in a standard Λ CDM model or any other particle with non-negligible thermal velocities. At the same time, exploring CDM through this framework might help understand more precisely the limitations of the N -body method by comparing results in the same way as various hydrodynamic solvers help understand the behaviour of the codes and their limits.

One could also argue (Sikivie 2010) that such an approach may be appropriate to simulate axions which remain quantum during the entire cosmological evolution. In such a case, the real value of \hbar and particle mass would have to be used, which would, however, probably lead to very high computational costs.

In this paper, we presented the validity of the method but many promising and interesting options have not yet been explored. The first obvious domain to investigate is the dependency on \hbar of the results. Early results tend to show that it may not be a crucial issue thanks to the universal gravitational profiles being low-degree power laws and hence generating only small quantum corrections to the Vlasov equation. It still remains an open question.

The other important area of investigation is the generation of initial conditions for more general cases than simple CDM. The procedures presented here cannot be applied without making some educated guess on the best shape of harmonic functions or without having to solve gigantic matrix eigenvalue problems. Combining some of these procedures or using interpolation techniques between lattice points are possible improvements worth exploring.

Finally, on the implementation side, lot of work can be done to make the codes more efficient. We already discussed the possibility of using an adaptive mesh to refine the grid in the regions of interest. It may also be possible to use an adaptive value of c and of the time step in the same way as N -body codes use different time bins for different particles. The locality of the interactions is an important feature as it makes the parallelization of the code straightforward. Running such a simulation on big clusters could thus be easily achieved without having to worry too much about complex communications and scalability issues.

Let us conclude by stating that our approach has a number of attractive features. Most importantly, the full phase space information is encoded in the wavefunctions. Working with many wavefunctions, we are in principle able to represent any given phase space distribution, including those where the velocity dispersion is important. Potentially, this would allow for numerical simulations of structure formation in presence of WDM.

ACKNOWLEDGEMENTS

This work was supported by the Swiss National Science Foundation and by the Tomalla Foundation. We would like to thank S. Cole, A. Maccio, J. Read and T. Theuns, for useful comments and discussions. OR acknowledges the support in part by the National Science Foundation under Grant no. PHYS-1066293 and the hospitality of the Aspen Center for Physics.

REFERENCES

- Abel T., Hahn O., Kaehler R., 2012, MNRAS, 427, 61
- Angulo R. E., Hahn O., Abel T., 2013, MNRAS, 434, 3337
- Aschbacher W. H., 2001, PhD thesis, Zürich
- Ballentine L., 1998, Quantum Mechanics: A Modern Development. World Scientific Press, Singapore
- Barnes J., Hut P., 1986, Nature, 324, 446
- Bastiaans M. J., 1997, in Mecklenbrücker W. F. G., Hlawatsch F., eds, Application of the Wigner Distribution Function in Optics. Elsevier, Amsterdam, p. 375
- Becker C., 2011, PhD thesis, SB, Lausanne
- Bernardeau F., Colomby S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
- Bertschinger E., 1995, NASA STI/Recon Technical Report N, 96, 22249

- Bertschinger E., 1998, *ARA&A*, 36, 599
- Binney J., Tremaine S., 2008, *Galactic Dynamics*, 2nd edn. Princeton Univ. Press, Princeton
- Boyarsky A., Ruchayskiy O., Shaposhnikov M., 2009, *Annual Review of Nuclear and Particle Science*, 59, 191
- Coles P., 2002, *MNRAS*, 330, 421
- Coles P., 2003, *The Wave Mechanics of Large-Scale Structure*. Kluwer, Dordrecht, p. 219
- Coles P., Spencer K., 2003, *MNRAS*, 342, 176
- Colín P., Avila-Reese V., Valenzuela O., 2000, *ApJ*, 542, 622
- Colín P., Valenzuela O., Avila-Reese V., 2008, *ApJ*, 673, 203
- Davies G., Widrow L. M., 1997, *ApJ*, 485, 484
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Dehnen W., 2000, *ApJ*, 536, L39
- Dehnen W., 2001, *MNRAS*, 324, 273
- Dehnen W., Read J. I., 2011, *Eur. Phys. J. Plus*, 126, 55
- Ercolessi E., Marmo G., Morandi G., Mukunda N., 2007, *J. Phys.: Conf. Ser.*, 87, 012010
- Fillmore J. A., Goldreich P., 1984, *ApJ*, 281, 1
- Fröhlich J., Graffi S., Schwarz S., 2007, *Commun. Math. Phys.*, 271, 681
- Fröhlich J., Tsai T.-P., Yau H.-T., 2010, in Alon N., Bourgain J., Connes A., Gromov M., Milman V., eds, *Visions in Mathematics*. Birkhäuser, Basel, p. 57
- Hillery M., O'Connell R., Scully M., Wigner E., 1984, *Phys. Rep.*, 106, 121
- Hirshfeld A. C., Henselder P., 2002, *Am. J. Phys.*, 70, 537
- Hockney R. W., Eastwood J. W., 1988, *Computer simulation using particles*. Hilger, Bristol
- Hudson R., 1974, *Rep. Math. Phys.*, 6, 249
- Husimi K., 1940, *Proc. Phys. Math. Soc. Jpn.*, 22, 264
- Jin S., Liao X., Yang X., 2008, *J. Hyperbolic Differ. Equ.*, 5, 569
- Joyce M., 2008, *Commun. Nonlinear Sci. Numer. Simul.*, 13, 100
- Klypin A. A., Trujillo-Gomez S., Primack J., 2011, *ApJ*, 740, 102
- Kravtsov A. V., 1999, PhD thesis, New Mexico State University
- Lee H.-W., 1995, *Phys. Rep.*, 259, 147
- Lions P.-L., Paul T., 1993, *Rev. Mat. Iberoamericana*, 9, 553
- Lovell M. R. et al., 2012, *MNRAS*, 420, 2318
- Ludlow A. D., Porciani C., 2011, *MNRAS*, 431, 1961
- Majda A. J., Majda G., Zheng Y., 1994, *Phys. D: Nonlinear Phenom.*, 74, 268
- Markovitch P. A., Mauser N. J., 1993, *Math. Models Methods Appl. Sci.*, 3, 109
- Mauser N. J., 2002, *J. Equ. Dérivées Partielles*, 11, 12
- Moyal J. E., 1949, *Math. Proc. Camb. Philos. Soc.*, 45, 99
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Peebles P. J. E., 1987, *ApJ*, 317, 576
- Plewa T., Linde T., Weirs V., 2005, in *Lecture Notes in Computational Science and Engineering*, Adaptive Mesh Refinement - Theory and Applications. Springer, Berlin
- Sakurai J. J., Napolitano J., 2011, *Modern Quantum Mechanics*. Addison-Wesley, Reading, MA
- Shandarin S., Habib S., Heitmann K., 2012, *Phys. Rev. D*, 85, 083005
- Short C. J., Coles P., 2006, *Journal of Cosmology and Astropart. Phys.*, 12, 12
- Sikivie P., 2010, *Int. J. Mod. Phys. A*, 25, 554
- Springel V. et al., 2005, *Nature*, 435, 629
- Teyssier R., 2002, *A&A*, 385, 337
- Tolman R., 1934, *Proc. Natl. Acad. Sci. USA*, 20, 169
- Wang J., White S. D. M., 2007, *MNRAS*, 380, 93
- Weinberg S., 1972, *Gravitation and cosmology: principles and applications of the general theory of relativity*. Wiley, New York
- Widrow L. M., Kaiser N., 1993, *ApJ*, 416, L71
- Wigner E., 1932, *Phys. Rev.*, 40, 749
- Woo T.-P., Chiueh T., 2009, *ApJ*, 697, 850
- Yoshikawa K., Yoshida N., Umemura M., 2013, *ApJ*, 762, 116
- Zhang P., Zheng Y., Mauser N. J., 2002, *Commun. Pure Appl. Math.*, 55, 582

APPENDIX A: SPHERICALLY SYMMETRIC CASE

The framework presented in Section 4 can be simplified in the case of (spatially) spherically symmetric distribution functions. The dimensionality of the problem is then reduced and allows more comprehensive convergence studies thanks to the lower number of discretization points needed.

If we consider only radial motion, then the distribution function can only depend on the distance to the centre r , the radial velocity v_r and the angle between those two vectors. We choose to use the cosine of this angle as our coordinate, denoted as y in what follows. The gravitational potential does only depend on the distance to the centre. We thus have $f \equiv f(r, v_r, y)$ and $U \equiv U(r)$. The density at a given r and total mass can be expressed using these new coordinates and read as

$$\rho(r) = \frac{2\pi}{a^3(\tau)} \int_0^\infty v_r^2 dp_r \int_{-1}^1 dy f(r, v_r, y), M \quad (A1)$$

$$= 4\pi \int_0^\infty r^2 dr \rho(r). \quad (A2)$$

It can be shown that the total mass is a conserved quantity under the equations of motion for f . The V-P system using those coordinates and assuming spherical symmetry becomes

$$\begin{aligned} \frac{\partial f}{\partial \tau} + \frac{y v_r}{a(\tau)} \frac{\partial f}{\partial r} - a(\tau) \frac{\partial U}{\partial r} \left[y \frac{\partial f}{\partial v_r} + \frac{(1-y^2)}{v_r} \frac{\partial f}{\partial y} \right] \\ + \frac{(1-y^2) v_r}{ra(\tau)} \frac{\partial f}{\partial y} = 0, \end{aligned} \quad (A3)$$

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial U}{\partial r} \right) = 4\pi G a(\tau)^2 (\rho(r) - \bar{\rho}). \quad (A4)$$

It may, in principle, be possible to find a Wigner-like distribution function for which the Wigner equation corresponds to this Vlasov equation. The wavefunctions entering such a distribution would probably obey a spherically symmetric version of Schrödinger's equation. This is, however, not the only way to handle this system.

The distribution function can be decomposed in two parts, one for each sign of the coordinate y :

$$f(r, v_r, y) = f_-(r, v_r) \delta_-(y+1) + f_+(r, v_r, \tau) \delta_+(y-1), \quad (A5)$$

where $\delta_\pm(x)$ are Dirac distributions defined on the interval $[-1, 1]$ only. We can then integrate over y and obtain two equations, one for f_+ and another identical up to the signs for f_- together with a boundary condition ensuring that the two distributions match when they reach $r = 0$ or $v_r = 0$. The next step in the procedure is to rescale these distribution functions by introducing $g_\pm(r, v_r) = f_\pm(r, v_r) r^2 v_r^2$ and define a combined distribution $h(r, v_r)$ such that

$$h(r, v_r) = \begin{cases} g_+(|r|, |v_r|) & \text{if } r v_r > 0 \\ g_-(-|r|, |v_r|) & \text{if } r v_r < 0. \end{cases} \quad (A6)$$

This new distribution function will obey the following Vlasov equation:

$$\frac{\partial h}{\partial \tau} + \frac{v_r}{a(\tau)} \frac{\partial h}{\partial r} - a(\tau) \frac{\partial U}{\partial r} \frac{\partial h}{\partial v_r} = 0, \quad (A7)$$

which is identical to the one-dimensional Vlasov equation (3). The difference being in the definition of density and mass that now reads

$$\rho(r) = \frac{2\pi}{r^2 R^3(\tau)} \int_{-\infty}^\infty dv_r h(r, v_r, \eta), M \quad (A8)$$

$$M = \frac{4\pi^2}{a^3(\tau)} \int_{-\infty}^{\infty} dr \int_{-\infty}^{\infty} h(r, v_r, \eta) db. \quad (\text{A9})$$

As we are back to the well-known case of Cartesian coordinates (at least for the Vlasov equation), we can introduce the same decomposition in terms of wavefunctions than in Section 4.2. We will thus solve a set of one-dimensional Cartesian Schrödinger equations alongside a three-dimensional spherically symmetric Poisson equation with a slightly odd density definition. Using the usual trick $V(r) = U(r)ra(\tau)$, the Laplacian term in Poisson's equation can be simplified and the system we want to evolve reads

$$i\hbar \frac{\partial \psi_n}{\partial t} = -\frac{\hbar^2}{2a(\tau)} \frac{\partial \psi_n^2}{\partial r^2} + m \frac{V}{r} \psi_n, M \quad (\text{A10})$$

$$\frac{\partial^2 V}{\partial r^2} = 4\pi G r \left(\frac{2\pi}{r^2} \sum_n \lambda_n |\psi_n(r)|^2 - \frac{4\pi^2 \Xi}{V_{\text{tot}}} \right), \quad (\text{A11})$$

where Ξ is the normalization of the wavefunctions that can be related to the total mass of the system through

$$M = \frac{4\pi^2}{a^3(\tau)} \sum_n \lambda_n \int_{-\infty}^{\infty} |\psi_n(r)|^2 dr = \frac{4\pi^2 \Xi}{a^3(\tau)}. \quad (\text{A12})$$

A dynamical term can then be added to equation (A11) to make the framework entirely local as discussed in Section 4.5. The system can eventually be evolved as if it was a purely one-dimensional problem. The only difference being the more complicated density terms sourcing Klein–Gordon's (or Poisson's) equation and the $1/r$ term in the potential of Schrödinger's equation.

The generation of initial conditions can be done in exactly the same way as outlined in Section 5. The only difference being the use of the modified distribution $h(r, v_r)$ (equation A6) instead of $f(r, v_r, y)$ as the starting point of the procedure.

This paper has been typeset from a \LaTeX file prepared by the author.